

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12Q 1/68, C12P 19/34, C07H 21/04, A61K 37/70</b>	<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 95/34679</b> <b>(43) International Publication Date:</b> 21 December 1995 (21.12.95)									
<b>(21) International Application Number:</b> PCT/US95/07605 <b>(22) International Filing Date:</b> 16 June 1995 (16.06.95)  <b>(30) Priority Data:</b> <table border="0"><tr><td>9412054.0</td><td>16 June 1994 (16.06.94)</td><td>GB</td></tr><tr><td>9502728.0</td><td>13 February 1995 (13.02.95)</td><td>GB</td></tr><tr><td>9507640.2</td><td>12 April 1995 (12.04.95)</td><td>GB</td></tr></table> <b>(71) Applicant (for all designated States except US):</b> THE GOVERNMENT OF THE UNITED STATES OF AMERICA, represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Office of Technology Transfer, National Institutes of Health, Suite 325, 6011 Executive Boulevard, Rockville, MD 20852 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> GONZALEZ, Frank, J. [US/US]; Unit 101, 5000 Battery Lane, Bethesda, MD 20814 (US). IDLE, Jeffrey, R. [GB/GB]; 19 Leazes Crescent, Newcastle, Upon Tyne NE1 4LW (GB).  <b>(74) Agents:</b> FEILER, William, S. et al.; Morgan & Finnegan, L.L.P., 345 Park Avenue, New York, NY 10154 (US).		9412054.0	16 June 1994 (16.06.94)	GB	9502728.0	13 February 1995 (13.02.95)	GB	9507640.2	12 April 1995 (12.04.95)	GB	<b>(81) Designated States:</b> AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).  <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>
9412054.0	16 June 1994 (16.06.94)	GB									
9502728.0	13 February 1995 (13.02.95)	GB									
9507640.2	12 April 1995 (12.04.95)	GB									
<b>(54) Title:</b> DEFECTS IN DRUG METABOLISM											
<b>(57) Abstract</b> <p>The invention relates to genetic material, and specifically portions of DNA, for identifying the presence or absence of a mutation in the drug metabolism gene CYP2C9 and CYP2A6. Further, the invention comprises a method for determining such mutations and a kit incorporating the genetic material of the invention for performing the said methods so as to determine the presence or absence of mutations in the drug metabolizing gene CYP2C9 and CYP2A6.</p>											

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

- 1 -

DEFECTS IN DRUG METABOLISMFIELD OF THE INVENTION

The invention relates to genetic material, specifically primers, for use in a method designed to determine the genotype of an individual; and also a kit, including the genetic material of the invention, for performing the method of the invention.

BACKGROUND OF THE INVENTION

It is well known that genetic polymorphisms in drug metabolizing genes give rise to a variety of phenotypes. This information has been used to advantage in the past for developing genetic assays that predict phenotype and thus predict an individual's ability to metabolize a given drug. The information is of particular value in determining the likely side effects and therapeutic failures of various drugs. The availability of this sort of information will result in routine phenotyping being recommended for certain categories of patients.

Drug metabolism is carried out by the cytochrome P450 family of enzymes. For example, the cytochrome P450 isozyme gene, CYP2C9 encodes a high affinity hepatic [S]-warfarin 7-hydroxylase which appears to be principally responsible for the metabolic clearance of the most potent enantiomer of warfarin. Similarly, the cytochrome P450 isozyme gene, CYP2A6, encodes a protein that metabolizes nicotine and coumarin and activates the tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone) (NNK).

It is of note that the above gene products are also known to metabolize other substrates, for example, the CYP2C9 gene product is also known to metabolize Tolbutamide, Phenytoin, Ibuprofen, Naproxen, Tienilic acid, Diclofenac and Tetrahydrocannabinol.

It follows that genetic polymorphisms or

- 2 -

0 mutations in either of the two aforementioned genes can lead to an impairment in metabolism of at least the aforementioned drugs.

5 In so far as CYP2C9 is concerned, sequences reported by Yasumori et al (1987 *J. Biochem.* 102:1075-1082.) and Kimura et al (1987 *Nuc. Acids Res.* 15:10053-10054) show differences at several positions including a C to T base change that results in a Arginine/Cysteine polymorphism at amino acid 144. This polymorphism has been designated R144C.

10 In so far as CYP2A6 is concerned, a T to A base change at position 488 of the cDNA sequence described by Yamano et al (1990 *Biochemistry* 29:1322-1329) results in substitution of Leucine 160 by Histidine. Henceforth this mutant form of the gene will be designated CYP2A6v1.

15 The variant CYP2A6v1 encodes an enzyme that is unstable and catalytically inactive. It is found in the general population at a frequency of about 1% but does not account for all slow metabolizers of coumarin.

20 Since the cDNA sequence structure of CYP2C9 and CYP2A6 are known, and since it is also known to perform genetic assays to determine whether a preselected mutation is present within a given gene, it should, in theory, be possible to design assays which specifically determine whether either of the aforementioned mutations are present  
25 in each of the respective aforementioned genes.

30 However, we have found an extraordinarily high degree of exon homology in the cytochrome P450 genes. This has resulted in non-specific binding of assay materials and poor performance of assays. In the instance where primers have been used to hybridize to genetic material, non-specific binding of such primers has taken place, and in the further instance where primers have been used to hybridize to genetic material with a view to performing a polymerase chain reactions we have found that  
35 related genes have also been amplified, for example, CYP2A7, CYP2A12 and CYP2C8 have also been amplified.

- 3 -

SUMMARY OF THE INVENTION

The present invention relates to novel variant alleles in cytochrome P450 genes which express enzymes involved in the metabolism of particular drugs and/or chemical carcinogens.

One object of the present invention relates to the discovery of new mutant or variant CYP2A6 alleles wherein the human gene is characterized. A new variant allele has been found which is designated CYP2A6v2. The cDNA and genomic sequence of CYP2A6v2 is provided in the present invention. Another new gene related to CYP2A6 has been discovered and is designated CYP2A13. The cDNA and genomic sequence of CYP2A13 is provided in the present invention.

Another object of the present invention relates to the use of intron sequences to specifically identify CYP2A6 and CYP2C9 variants in a gene specific detection assay.

Another object of the present invention is to use an oligonucleotide probe, specific for regions unique to a particular CYP2 variant to screen for the presence or absence of the variant in a sample.

Yet another object of the invention is to provide genetic material, a method, and a kit which enable genotyping of the CYP2C9 and CYP2A6 gene with a view to providing phenotypic information concerning drug metabolism.

A further object of the present invention provides a method for diagnostically determining the sensitivity of a patient for specific drugs and chemical carcinogens. Such a method is widely applicable in determining the proper dosage of a drug for a patient.

Another object of the present invention provides a method of genotyping CYP2A6 and CYP2C9 and determining whether a mutation has altered the sequence of these genes and hence altered sensitivity to particular drugs and chemical carcinogens.

- 4 -

In accordance with the present invention a method is provided which utilizes the finding that each variant of a CYP2 gene has specific nucleotide differences as compared with the wild-type CYP2 gene. Such nucleotide changes can be utilized in a probe-hybridization assay, which is capable of specifically detecting a chosen variant and not other variants.

The present invention also provides a genotyping method for identifying the presence or absence of a mutation at codon 144 of the coding sequence of CYP2C9, or alternatively, at codon 160 of the coding sequence of CYP2A6, or alternatively, a gene conversion event involving CYP2A6 and CYP2A7 in exons 3, 6 or 8 comprising use of a portion of DNA. Such a mutation is then correlated to the sensitivity of particular drugs and chemical carcinogens.

The present invention further relates to a gene-specific bioassay which is capable of distinguishing between the CYP2 genes and identify the presence or absence of a mutation in CYP2A6 and CYP2C9 genes. Such a bioassay can diagnostically predict the sensitivity of an individual to particular drugs or chemical carcinogens. For example, the CYP2C9 variants identify a sensitivity to a commonly used anti-coagulant drug, warfarin. The CYP2A6 variants identify sensitivity to coumarin, nicotine and nitrosamines. The sensitivity to nicotine may be used to predict a predisposition to tobacco-related diseases, a propensity to smoking and adverse reactions to exposure to nicotine. Further, CYP2A6 genes are associated with the activation of nitrosamines, elevated levels of which have been correlated with many cancers.

The present invention also provides a method of genotyping the CYP2A6 and CYP2C9 genes using allele-specific amplification reaction.

In addition, a highly-specific combination genotyping bioassay has been developed to identify mutations within CYP2A6 and CYP2C9 which are linked to

- 5 -

sensitivity to particular drugs and chemical carcinogens. This combination bioassay comprises a gene-specific amplification reaction, an exon-specific amplification reaction and an endonuclease cleavage reaction wherein only one form, either mutant or wild-type is cleaved, producing either a single nucleic acid fragment or multiply nucleic acid fragments depending upon the presence or absence of the mutation. For example, one CYP2C9 variant, R144C, which contains a C<sub>472</sub>→T mutation can be identified by an *Ava*II restriction site. CYP2A6 variants can also be identified by their corresponding mutations. CYP2A6v1 which contains a T<sub>488</sub>→A mutation can be identified by a *Xcm*I restriction site. CYP2A6v2 which contains a T<sub>415</sub>→A mutation can be identified by a *Dde*I restriction site.

The present invention also relates to a method for screening patients for drug sensitivity prior to their treatment with that drug, thereby alerting a physician of a drug sensitivity. In addition, the method may be used to screen patients for a predisposition to cancers related to excessive nitrosamine activation, which are associated with mutations within the CYP2A6 gene locus. Further, the method may be used to screen patients for a sensitivity to chemical carcinogens, based upon the genotype of the CYP2A6 and/or CYP2C9 alleles.

One such new allele variant, CYP2A6v2, has 98% nucleotide similarity and 80% amino acid similarity with the wild type CYP2A6, respectively. The present invention relates to the new CYP2A6v2 variant, the cDNA sequence and its genomic sequence wherein the alterations in sequence are within exons 3, 6 and 8, which are attributed to a gene conversion. In addition, another new gene, also involved in drug metabolism has been identified, and has been designated CYP2A13. This gene plays a similar role in drug metabolism as CYP2A6. These new gene sequences or fragments thereof are used as probes in identifying specific CYP2 variants in samples. In additions,

- 6 -

fragments of the new genes are used as primers in a genotyping assay.

The invention further provides isolated CYP2Av2 and CYP2A13 cDNAs for use in gene therapy and replacement protocols for individuals who are predisposed to sensitivity to needed drugs or to chemical or environmental carcinogens.

In accordance with an aspect of the present invention, there are provided primary human cells which are genetically engineered with CYP2A6v2 or CYP2A13 DNA (RNA) which encodes a therapeutic agent of interest, and the genetically engineered cells are employed as a therapeutic agent. (The term "therapeutic," as used herein, includes treatment and/or prophylaxis.)

Gene expression in an organism in accordance with the practices of this invention is regulated, inhibited and/or controlled by incorporating in or along with the genetic material of the organism non-native DNA which transcribes to produce an RNA which is complementary to and capable of binding or hybridizing to a mRNA produced by a gene located within said organism. Upon binding to or hybridization with the mRNA, the translation of the mRNA is prevented. Consequently, the protein coded for by the mRNA is not produced. In the instance where the mRNA translated product, e.g. protein, is vital to the growth of the organism or cellular material, the organism is so transformed or altered such that it becomes, at least, disabled.

Accordingly, in the practices of this invention from a genetic point of view as evidenced by gene expression, new organisms are readily produced. Further, the practices of this invention provide a powerful tool or technique for altering gene expression or organisms through gene therapy. The practices of this invention may cause the organisms to be disabled or incapable of functioning normally or may impart special properties thereto. The DNA of CYP2A6v2 or CYP2A13 employed in the

- 7 -

practices of this invention can be incorporated into the treated or effected organisms by direct introduction into the nucleus of a eukaryotic organism or by way of a plasmid or suitable vector containing the special DNA of this invention in the case of a procaryotic organism.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the invention are described by way of example only with reference to the accompanying figures wherein:

Fig. 1. Shows the sequence of exon 2, intron 2 and exon 3 of CYP2C8 and CYP2C9, cDNA sequences (from 4) are shown at the top of the page together with sequences from 6 genomic clones encompassing exon 2, intron 2 and exon 3 of CYP2C8 and CYP2C9. The position of the polymorphism at codon 144 of CYP2C9 and the PCR primers are indicated.

Fig. 2. Shows the sequence of intron 2, exon 3 and intron 3 of CYP2A6, CYP2A7 and CYP2A12. The position of the polymorphism at codon 160 in CYP2A6 and the PCR primers are indicated.

Fig. 3. Shows the detection of CYP2C9 Arg<sub>144</sub> Cys polymorphism by PCR. Following amplification, samples were digested with AvaII and analyzed on a 1.8 % agarose gel. Lane 1 and lanes 3 to 6 show homozygous wild-type subjects, lane 2 a heterozygous individual and lane 7 undigested PCR product.

Fig. 4. Shows detection of CYP2A6 Leu<sub>160</sub>. His polymorphism by PCR. Two parallel PCR reactions were carried out and the products analyzed on a 1 % agarose gel. Lanes 1, 3, 5 and 7 show the results of the wild-type specific assay and lanes 2, 4, 6 and 8 the results of the variant-specific assay for the same four subjects. Subjects 1 and 2 (lanes 1-4) are homozygous wild-type, subject 3 (lanes 5 and 6) heterozygous and subject 4 (lanes 7 and 8) homozygous for the mutation.

- 8 -

Fig. 5. Shows distribution of the weekly maintenance doses for warfarin in patients (n=57) homozygous for the CYP2C9 wild-type allele (open bars) and heterozygous (n=37) for the R144C mutant allele (solid bars). Arrows show the median weekly dose requirement of warfarin for each genotype.

Fig. 6. Represents 7-hydroxylation of coumarin (%) in a family genotyped for the CYP2A6 and CYP2A6v1 alleles, showing a subject homozygous for the CYP2A6v1 allele who is deficient in coumarin 7-hydroxylation.

Fig. 7. Shows the difference between the genomic and cDNA sequences for the CYP2A6 gene.

Figs. 8a and b. Shows the conversion event which leads to the CYP2A6v2 allele.

Figs. 9a through 9c. Shows the detection of CYP2A6v2 by PCR. (Fig. 9A) gene-specific amplification by PCR of the CYP2A6 gene using E3F and E3R. Lanes 1 to 4 show the 7.8 Kb band obtained from several representative human genomic DNA templates, lane 5 correspond to a negative control in the absence of template and lane 6 contains 1 Kb DNA ladder (GIBCO BRL) as six markers. (Fig. 9B) Exon-specific PCR amplification of exon 3 from the 7.8 Kb long-PCR product and restriction endonuclease pattern obtained after digestion with *XcmI* (left) and *DdeI* (right) to detect the CYP2A6v1 and CYP2A6v2 alleles, respectively. The genotypes shown correspond to: wild type (+/+), heterozygous (+/-) and homozygous (-/-) subjects. (C) The genotyping strategy which has been developed. Exons are indicated by boxes. The position of the corresponding primer pairs are indicated by horizontal arrows. *XcmI* and *DdeI* restriction sites generate digestion patterns for the different alleles having fragment sizes as shown.

Fig. 10. Schematic diagram depicting methodology underlying a CYP2C9 genotyping assay.

Fig. 11. CYP2A6v2 cDNA sequence.

Fig. 12. CYP2A6v2 genomic DNA sequence having

- 9 -

7216 base pairs.

Fig. 13. CYP2A13 cDNA sequence.

Fig. 14. CYP2A13 genomic DNA sequence having 8779 base pairs.

Fig. 15. Agarose minigel electrophoresis of PCR products. The CYP2C9 wild-type allele (Arg-144) and R144C respectively, Lanes marked "+/+" and "+/-" contain homozygous wild types and heterozygotes respectively. the right-hand lane contains a 100 bp ladder.

#### DETAILED DESCRIPTION OF THE INVENTION

The cytochrome P450 isozyme gene, CYP2C9 encodes a high affinity hepatic [S]-warfarin 7-hydroxylase which appears to be principally responsible for the metabolic clearance of the most potent enantiomer of warfarin along with metabolizing a number of other drugs and chemical carcinogens. Similarly, the cytochrome P450 isozyme gene, CYP2A6, encodes a protein that metabolizes nicotine, coumarin and a host of other drugs and chemical carcinogens CYP2A6 also activates the tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (herein referred to as "NNK"). Many cancers have been associated with activation and/or accumulation of nitrosamines. The present invention allows detection of a predisposition to such cancers.

It is of note that the above gene products are also known to metabolize other substrates. For example, the CYP2C9 gene product is also known to metabolize Tolbutamide, Phenytoin, Ibuprofen, Imipramine, Naproxen, Tienilic acid, Diclofenac and Tetrahydrocannabinol and hence can also be used to detect sensitivities to these drugs. A list of CYP2C9 drug substrates has been documented and is incorporated herein by reference (Gonzalez & Idle 1994 *Clin. Pharmacokinet* 26:59-70). Hence, the present invention can be used to screen for sensitivities to these drugs.

In addition, CYP2C9 has been associated with the metabolism of chemical carcinogens, such as polycyclic

- 10 -

aromatic hydrocarbons. For example, the most ubiquitous environmental carcinogen, benz-[a]-pyrene is metabolized by CYP2C9. Benz-[a]-pyrene is found in tobacco, barbecued meats, car exhaust and generally, in polluted air. This compound, as it accumulates in the body becomes a potent DNA intercalating agent, ultimately resulting in cell transformation and the formation of tumors. The present invention provides a diagnostic method of screening individuals for their ability to metabolize and hence inactivate benz-[a]-pyrene. For example, a homozygote wild-type CYP2C9 individual would be better able to tolerate high levels of benz-[a]-pyrene than a heterozygote of the CYP2C9 allele.

Similarly, the CYP2A6 allele is associated with drug sensitivity and carcinogen metabolism. Coumarin sensitivity is directly related to the presence of a variant CYP2A6 allele, such as CYP2A6v1, CYP2A6v2 and also CYP2A13. Coumarin is a drug used in treatment of neoplastic diseases, such as lymphomas. (See Martindale: The Extra Pharmacopoeia 1993 Ed. Reynolds, J.E.F., The Pharmaceutical Press, London, p. 1358). Its suggested dosage is very high. Therefore, the present invention is useful in determining a patient's sensitivity to the drug in order to prescribe a proper dosage and avoid toxicity.

Another drug, Thiotepe™, is used in the treatment of a variety of neoplastic diseases, such as in treating women with breast cancer and children with brain tumors. Thiotepe is metabolized by CYP2A6 into Tepa, which is an intermediate more therapeutically potent than Thiotepe. Therefore, if a patient has a very active CYP2A6 enzyme, it is likely the patient will require lower doses of Thiotepe to provide a therapeutically effective amount. As one can see, the dosage provided to a patient is dependent upon the rate a patient is capable of metabolizing activating the drug. The present invention has identified variant alleles whose enzymatic activity is compromised. In addition, the present invention provides

- 11 -

° a simple method of genotyping patients for Thiotepa drug sensitivity. With information concerning patient sensitivity to such drugs, the proper dosage can be provided, hence maximizing drug efficiency and minimizing drug toxicity.

5 Further, CYP2A6 has been associated with nicotine metabolism. In addition to being an active ingredient in tobacco, nicotine also has several clinical uses. Nicotine is used clinically to treat various neurological disorders, such as Parkinson's disease and  
10 Alzheimer's disease. In addition, nicotine is used to treat tobacco addiction. In all of these situations, it is important to know a patient's sensitivity to nicotine, since extremely sensitive patients will become violently ill upon administration of nicotine. Therefore the  
15 present invention provides a method of identifying nicotine-sensitive patients by genotyping a patient's CYP2A6 allele. The present invention also provides a convenient method for determining an individual's general predisposition to using tobacco based upon their  
20 sensitivity to nicotine.

In addition, CYP2A6 is involved in activating nitrosamines, thereby producing the potent carcinogen NNK. Increased levels of NNK have been associated with a variety of cancers, including but not limited to lung  
25 cancer, nasal-pharynx cancers, throat cancers and colon cancers. In general, elevated levels of CYP2A6 has been associated with cancers associated with exposure to nitrosamines. The present invention may detect a patient's predisposition to such cancers. The presence of  
30 a CYP2A6 gene or a variant thereof will affect the likelihood that procarcinogens present in tobacco smoke will be activated into carcinogenic nitrosamines and nitrosamine-derivatives and therefore result in the development of a cancer.

35 It follows that genetic polymorphisms or mutations in either of the two aforementioned genes can

- 12 -

lead to an impairment in metabolism of at least the  
aforementioned drugs and chemical carcinogens.

The present invention relates to the  
identification of the absence or presence of mutations in  
CYP2C9 and CYP2A6 and thus predict the phenotype of an  
individual and so predict whether and how an individual is  
likely to metabolize particular drugs and chemical  
carcinogens. For instance, the R144C mutation arising  
from a C<sub>472</sub>→T base substitution in the CYP2C9 gene results  
in a reduction in warfarin metabolism. This implies that  
patients with this mutation receiving warfarin require a  
lower dose to maintain an anticoagulation target than  
those patients who do not have the mutation and are also  
receiving warfarin. Conversely, homozygous wild-types  
require higher doses in order to maintain an  
anticoagulation target.

"Mutation", as the term is used herein denotes  
an allelic variation of a known sequence, which alters the  
expressed gene product's activity. Such a variation need  
not completely inactivate the gene product's activity but  
merely alter it.

Similarly, one mutation within CYP2A6v1 arising  
from a T<sub>488</sub>→A base change results in substitution of  
Leucine 160 by Histidine. Another CYP2A6 variant,  
CYP2A6v2, has been identified which differs from CYP2A6 in  
the regions of exons 3, 6 and 8. One particular mutation  
in CYP2A6v2, T<sub>415</sub>→A mutation is useful in the assay of the  
present invention. These substitutions are very useful in  
detecting predispositions to cancers associated with  
tobacco and activation of nitrosamines. The normal CYP2A6  
enzyme functions in the metabolism of nicotine, one of the  
carcinogenic compounds in tobacco.

In addition, the present invention relates to  
the identification of a new variant of CYP2A6 designated  
CYP2A6v2. The variations of CYP2A6v2 from CYP2A6 bear  
sequence relatedness with the corresponding exons of the  
CYP2A7 gene, suggesting a recent gene conversion. The cDNA

- 13 -

and genomic sequence for this gene is provided in the present invention. Hence, at least three different allelic variants of CYP2A6 exist and are illustrated in Figure 8. These allelic variants include CYP2A6, CYP2A6v1 and CYP2A6v2.

Further, the present invention relates to a new CYP2A gene, designated CYP2A13. This gene produces an inactive form of CYP2A6, however variants at particular positions, including amino acid positions 117, 209 and 365 produce an enzyme which may alter the enzyme's activity and hence affect drug sensitivity. These mutations in CYP2A6 are likely to result in a deficiency or impaired activity of one of the enzymes responsible, for example, for metabolizing drugs, nicotine and nitrosamines.

CYP2A13 is considered a new cytochrome P450 gene. However, since the CYP2A13 gene product has a similar function as the CYP2A6, it is discussed herein as a variant of CYP2A6. That is, assays using the specific mutated amino acid positions 117, 209 and 365 of CYP2A13 and detecting variations at those positions are indicative of CYP2A6-like variant functions.

In one embodiment, the CYP2A6v2 or CYP2A13 proteins or functional portions thereof are expressed as recombinant genes in a cell, so that the cells may be transplanted into an individual in need of gene therapy due to the predisposition to a carcinogen-associated cancer or a sensitivity to a drug. To provide gene therapy to an individual, a genetic sequence which encodes for all or part of the CYP2A6v2 or CYP2A13 ligands are inserted into vectors and introduced into host cells. Examples of vectors that may be used in gene therapy include, but are not limited to, defective retroviral, adenoviral, or other viral vectors (see, e.g., Mulligan, R.C., 1993, Science, 260:926-932). The means by which the vector carrying the gene may be introduced into the cell includes, but is not limited to, microinjection, electroporation, transduction, or transfection using DEAE-

- 14 -

o dextran, lipofection, calcium phosphate or other procedures known to the skilled routineer (see, e.g., Sambrook et. al. (Eds.), 1989, In "Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Press, Plainview, New York). Examples of cells into which the vector  
5 carrying the gene may be introduced include, but are not limited to, continuous culture cells, such as COS, NIH/3T3, and primary or culture cells of the relevant tissue type.

10 More specifically, there is provided a method of enhancing the therapeutic effects of blood cells, that are infused in a patient, comprising: (i) inserting into the blood cells of a patient a DNA (RNA) segment encoding CYP2A6v2 or CYP2A13 gene product that enhances the therapeutic effects of the blood cells; and (ii)  
15 introducing cells resulting from step (i) into the patient under conditions such that the cells resulting from step (i) "target" to a tissue site. In the alternative, as previously described the cells are not "targeted" and functions as a systemic therapeutic. The genes are  
20 inserted in such a manner that the patient's transformed blood cell will produce the agent in the patient's body. In the case of antigen-specific blood cells which are specific for an antigen present at the tissue site, the specificity of the blood cells for the antigen is not lost  
25 when the cell produces the product.

Alternatively, as hereinabove indicated, CYP2A6v2 or CYP2A13 DNA (RNA) may be inserted into the blood cells of a patient, in vivo, by administering such DNA (RNA) in a vehicle which targets such blood cells.

30 Further details regarding methods of gene therapy are provided in Anderson et al., U.S. Patent No. 5,399,343 which is herewith incorporated herein by reference.

35 In another embodiment, antisense CYP2A6v2 or CYP2A13 DNA or RNA may be used to control the expression of CYP2 gene. For example, antisense therapy may be used

- 15 -

to control CYP2A6's ability to activate dangerous nitrosamines by curbing its expression. Methods of producing such antisense molecules are described in U.S. Patent No. 5,190,931, which is incorporated herein by reference.

Developing a genotyping assay, which could distinguish the CYP2 genes of interest from other cytochrome P450 genes required careful engineering since these genes have a high degree of sequence homology. To overcome this problem, one embodiment of the present invention has elucidated the genomic sequence structure of CYP2C9 and CYP2A6 with a view to making, in part, intron specific primers. That is to say primers which, in part, hybridize to at least one intron, preferably an intron adjacent to an exon including the mutation of interest, in the gene to be examined. Since there is less homology between the introns of cytochrome P450 genes, it has been found that using intron specific primers, gene specific assay can be undertaken. The present invention has a further advantage of using intron specific primers in so far as the use of such primers facilitates the manufacture of an optimum length of DNA which in turn facilitates the specificity of the instant bioassay.

A "genotyping" assay as the term is used herein refers to any diagnostic or predictive test to detect the presence or absence of allelic variants of a known gene sequence at a specified gene locus. Two gene loci are of particular interest in the present invention, CYP2A6 and CYP2C9.

Further, the present invention relates to differences between the genomic DNA sequence structure and the cDNA sequence structure, as illustrated in Figure 7. As a result, primers directed at the genomic sequence structure have been developed which are more reliable.

Several methods are provided for identifying the presence or absence of a mutation at codon 144 of the coding sequence of CYP2C9, or alternatively, at codon 160

- 16 -

of the coding sequence of CYP2A6, or alternatively, a gene conversion event involving CYP2A6 and CYP2A7 in exons 3, 6 or 8 comprising a DNA encompassing the region of a CYP2 gene unique to that variant.

One such method relates to an assay which contemplates the use of one specific primer which specifically encompasses the region containing the mutation, and a second primer which is complementary to another portion of the gene. The second primer sequence chosen is based upon the CYP2A6, CYP2C9 or CYP2A13 sequences as set forth in figures 12, 1 and 14, respectively, depending upon the preferred size of the amplification product. One skilled in the art will know how to select second primer based on the region of gene chosen for amplification. These primers need not be identical to a given sequence but must be sufficiently complementary to hybridize to the target region in a specific manner. In short, the primers are preferably at least substantially homologous to the nucleic acid sequence provided.

Nucleic acid sequences includes, but is not limited to, DNA, RNA or cDNA. Nucleic acid sequence as used herein refers to an isolated nucleic acid sequence. Substantially homologous as used herein refers to substantial correspondence between the nucleic acid primer sequence of as described herein and that of any other nucleic acid sequence. Substantially homologous means about 50-100% homologous homology, preferably by about 70-100% homology, and most preferably about 90-100% homology between the particular sequence discussed and that of any other nucleic acid sequence.

In the instant application, the term "primer" is further used to designate a molecule comprising at least three nucleotides, the exact length being determined by the requisite amount of DNA needed, under given reaction conditions, to bind to or interact with a test sample so as to identify the presence or absence of either of said

- 17 -

mutations. Preferably, the primer is usually between 15 and ideally about 20 to 50 oligonucleotides in length.

The primer is selected, or adapted, to be substantially complementary to a part of DNA which is adjacent to the region of at least one of the  
5 aforementioned mutations. Thus such a primer is able to hybridize with a part of DNA that contains a region in which the mutation of interest may be found. Although the primer may not reflect the exact sequence of the region in which the mutation is thought to occur, the more closely  
10 the primer is to this sequence, then the better the binding will be. Ideally, the more closely the sequence of the 3' end of the primer is to said region the better the binding or interaction will be.

An alternative method for using the sequence  
15 unique to a variant for detection relates to use of an oligonucleotide probe for specifically detecting the presence or absence of a CYP2 variant gene in a sample. This method comprises the steps of contacting the sample with a nucleic acid probe, allowing hybridization, forming  
20 a probe: CYP2 variant complex; washing excess probe from probe: CYP2 variant complex; and detecting probe: CYP2 variant complex, wherein a positive signal is an indication of the presence of the CYP2 variant in the sample.

25 The hybridization of the probe to sample nucleic acids can be carried out by any of the methods commonly used in the art. Such methods include but are not limited to, Dot blot, Colony hybridization, Southern blot, solution hybridization and *in situ* hybridization.

30 Washing the excess probe from the probe: CYP2 variant DNA can be accomplished by many well-known methods. Simply rinsing the complex with excess buffer will facilitate removal of excess probe. Alternatively, washing may entail separating the probe: CYP2 variant  
35 complex from excess probe. Many methods are known to one skilled in the art and include but are not limited to

- 18 -

centrifugation, filtration and magnetic force.

According to the present invention there is provided a portion of DNA suitable for use as a primer in a method for identifying the presence or absence of a mutation either at codon 144 of the coding sequence of the gene CYP2C9, or alternatively, at least one gene conversion event involving CYP2A6 and CYP2A7 in exons 3, 6 or 8, or alternatively, at codon 160 of the coding sequence of the gene CYP2A6; comprising a DNA which is adapted to hybridize to at least one intron of at least one of said genes.

In one embodiment, the method comprises the use of at least one restriction endonuclease to digest DNA from individuals to be tested. In this instance, DNA from individuals positive for the wild-type form of CYP2C9 provide a digest with a restriction endonuclease, such as *Ava*II results in production of two fragments, a first fragment including 270 base pairs and a second fragment including 50 base pairs. In contrast, individuals having the aforementioned mutation in CYP2C9 present a single fragment of 320 base pairs only. This is due to a loss of the *Ava*II site. The CYP2A6 gene variants can also be distinguished by the occurrence of specific restriction endonuclease sites. The CYP2A6v1 variant, which is a T<sub>488</sub>→A mutation in exon 3 can be identified by a variant-specific *Xcm*I restriction site. The CYP2A6v2 variant, which contains a C<sub>415</sub>→A mutation within exon 3 can be identified by a variant-specific *Dde*I restriction site. The wild-type CYP2A6 gene does not contain either an *Xcm*I or *Dde*I site. The results of such restriction endonuclease digestions are illustrated in Figure 9.

It may be necessary to amplify the DNA prior to digestion. Such may be the case when the DNA of interest is present in minute quantities in a sample. In such circumstances, amplification of DNA to be tested is undertaken before digesting the DNA as described above. This provides for a greater quantity of materials.

- 19 -

Amplification is performed using any conventional technique, such as by a PCR reaction. Many other techniques for amplification can be used in producing sufficient DNA for detections. Such amplification techniques are well-known to the skilled artisan and include, but are not limited to polymerase chain reaction (PCR), PCR *in situ*, ligase amplification reaction (LAR), ligase hybridization, QB bacteriophage replicase, transcription-based amplification system (TAS), genomic amplification with transcript sequencing (GAWTS) and nucleic acid sequence-based amplification (NASBA). A general review of these methods is available in Landegren, et al., *Science* 242:229-237 (1988) and Lewis, R., *Genetic Engineering News* 10:1, 54-55 (1990), which is incorporated herein by reference.

One embodiment of the present invention uses oligonucleotide primers in an amplification and detection assay. A basic description of nucleic acid amplification is described in Mullis, U.S. Patent No. 4,683,202, which is incorporated herein by reference. The amplification reaction uses a template nucleic acid contained in a sample, two primer sequences and inducing agents. The extension product of one primer when hybridized to the second primer becomes a template for the production of a complementary extension product and vice versa, and the process is repeated as often as is necessary to produce a detectable amount of the sequence.

The inducing agent may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, *E.coli* DNA polymerase I, thermostable *Taq* DNA polymerase, Klenow fragment of *E.coli* DNA polymerase I, T4 DNA polymerase, other available DNA polymerases, reverse transcriptase and other enzymes which will facilitate combination of the nucleotides in the proper manner to form amplification products.

- 20 -

°  
A sample being screened for the presence or absence of a mutation in CYP2A6 and/or CYP2C9 genes can be tested with the instant invention. The nucleic acid material can be in purified or nonpurified form, provided the sample contains the CYP2A6 and/or CYP2C9 genes. The sample may be derived from any tissue or bodily fluid, wherein the patient's DNA can be found. A clinically practical type of sample is a blood specimen which contains patient DNA and can conveniently be genotyped in the bioassay of the present invention.

10 The "primers", as the term is used in the present invention refers to an oligonucleotide, whether occurring naturally as in a purified restriction digest or produced synthetically, which is capable of acting as a point of initiation of synthesis when placed under  
15 conditions wherein synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, i.e. in the presence of nucleotides and an inducing agent such as DNA polymerase and at a suitable temperature and pH. The primers are preferably single  
20 stranded for maximum efficiency in amplification, but may alternatively be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare amplification products. Preferably, the primers are oligodeoxyribonucleotides. The primers  
25 must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent. The exact lengths of the primers will depend on many factors, including temperature, source of primer and use of the method. For diagnostic methods, the primers  
30 typically contain at least 10 or more nucleotides. The oligonucleotide primers may be prepared using any suitable method, such as, for example, the phosphotriester and phosphodiester methods (Narang, S.A., et al., *Meth. Enzymol.* 68:90 (1979); Brown E.L., et al., *Meth. Enzymol.*, 68:109 (1979)) or automated embodiments thereof. In one  
35 such automated embodiment diethylphosphoramidites are used

- 21 -

as starting materials and may be synthesized as described by Beaucage et al., *Tetrahedron Letters* 22:1859-1962 (1981). One method for synthesizing oligonucleotides on a modified solid support is described in U.S. Pat. No. 4,458,066. It is also possible to use a primer which has been isolated from a biological source (such as a restriction endonuclease digest).

In a genotyping bioassay of the present invention, one embodiment comprises a gene-specific amplification reaction, an exon-specific amplification reaction and a restriction endonuclease reaction. In such a reaction a suitable polynucleotide polymerase is used in the amplification reaction, many of which have already been described in the art. In addition, any appropriate restriction endonuclease which is designed to digest the DNA and so provide information concerning genotype may be used.

It may further be necessary to provide a label on the nucleic acid for detection. The nucleic acid can be DNA or RNA and made detectable by any of the many labeling techniques readily available and known to the skilled artisan. Such methods include, but are not limited to, radio-labelling, digoxigenin-labeling, and biotin-labeling. A well-known method of labeling DNA is <sup>32</sup>P using DNA polymerase, Klenow enzyme or polynucleotide kinase. In addition, there are known non-radioactive techniques for signal amplification including methods for attaching chemical moieties to pyrimidine and purine rings (Dale, R.N.K. et al. 1973 *Proc. Natl. Acad. Sci. USA*, 70:2238-2242; Heck, R.F. 1968 *S. Am. Chem. Soc.*, 90:5518-5523), methods which allow detection by chemiluminescence (Barton, S.K. et al. 1992 *J. Am. Chem. Soc.*, 114:8736-8740) and methods utilizing biotinylated nucleic acid probes (Johnson, T.K. et al. 1983 *Anal. Biochem.*, 133:125-131; Erickson, P.F. et al. 1982 *J. of Immunology Methods*, 51:241-249; Matthaei, F.S. et al 1986 *Anal. Biochem.*, 157:123-128) and methods which allow detection by

- 22 -

o

fluorescence using commercially available products. Non-radioactive labelling kits are also commercially available. Such a label can readily be incorporated into the nucleic acid during an amplification step. In the absence of an amplification step, a target nucleic acid can readily be chemically or enzymatically modified to carry a label. Additionally, it may be preferable to provide a labeled primer which may serve to incorporate a label into the nucleic acid target. Probes, as may be used in an embodiment of the invention may also be chemically or enzymatically labeled as described above.

In a preferred embodiment of the invention said DNA primer hybridizes to an intron adjacent said position of said mutation. Preferably said DNA is a primer with the 3'-end specific for the gene of interest. Preferably further still said DNA is single stranded. Preferably further still, in so far as the CYP2C9 mutation is concerned, said primers are as follows:

HF18: position 8 of intron 2 onwards of genomic sequence in forward orientation comprises  
5' TGCAAGTGCCTGTTTCAGCA 3'  
HF2R: position 505 onwards of cDNA sequence in reverse orientation comprises  
5' AGCCTTGGTTTTTCTCAACTC 3'.

It is of note that both these primers are designed to be specific for CYP2C9 and so do not amplify related genes such as CYP2C8, which notably also has an Arginine<sub>144</sub> present.

Preferably, in so far as CYP2A6 is concerned, three primers J51, J61 and B are used in two parallel allele-specific PCR reactions. These primers are as follows:

J51 comprises 5' GGCTTCCTCATCGACGCACT 3'  
(forward strand from position 479 of cDNA

- 23 -

sequence described as hIIA3 (Yamano, et al. 1990 *Biochem* 29:1322-29)).

J61 comprises 5' GGCTTCCTCATCGACGCACA 3' (forward strand from position 479 of cDNA sequence described as hIIA3v (Yamano, et al. 1990 *Biochem* 29:1322-29)).

Both J51 and J61 contain a substitution at position 18 of A for C to give improved specificity as suggested by Newton et al (1989 *Nuc. Acids Res.* 17:2503-2516).

Primer B comprises 5' AATTCCAGGAGGCAGGGCCT 3' (reverse orientation from position 125 of intron 3 of CYP2A6 (onwards). Designed so that only CYP2A6 and not CYP2A7 or CYP2A12 are amplified.

One method of genotyping CYP2A6 provides an allele-specific amplification reaction method is used. In this instance, DNA which is adapted to specifically hybridize to the wild-type or the mutant type of the gene is incubated with test DNA under reaction conditions and the resultant products are analyzed by electrophoresis and then visualized by staining with ethidium bromide. Individuals who are homozygous for the wild-type allele produce a reaction product with primer J51 only. Similarly, individuals who are homozygous for the mutation produce a reaction product with primer J61 only. Those individuals who are heterozygous produce a reaction product with both J51 and J61.

Alternatively, another method for genotyping CYP2A6 is provided in a specific amplification bioassay, which is achieved with primers F4 and R4 as follows:

The F4 primer (forward) comprises 5' CCCCTTATCCTCCCTTGCTGGCTGTGTCCCAAGCTAGGCAGGATT CATGGTGGGGCA 3', wherein a preferred fragment thereof further comprises 5' CCTCCCTTGCTGGCTGTGTCCCAAGCTAGGC 3'.

- 24 -

The R4 primer (reverse) comprises

5' GCCACCACGCCCCCTTCCTTTCCGCCATCCTGCCCCCAGTCTTAGC  
TGCGCCCCCTCTC 3', wherein a preferred fragment  
thereof further comprises

5' CGCCCCCTTCCTTTCCGCCATCCTGCCCCCAG 3'.

This method of CYP2A6 genotyping involves a  
first amplification reaction with F4 and R4 primers, which  
generates a DNA fragment approximately 7.8 kb in size.

This amplification step is facilitated by polymerases  
which are capable of transcribing long stretches of DNA.  
To distinguish the CYP2A6v1 and CYP2A6v2 variant alleles,  
an exon-specific amplification step is carried out using  
the 7.8 Kb DNA fragment as template DNA. This may be  
accomplished using the following primer pair:

The E3F primer (forward) comprises

5' CCTGATCGACTAGGCGTGGTATTCAGCAACGGGGAGCGGCCAAG  
CAGCTCCTG 3', wherein a preferred fragment  
thereof further comprises

5' GCGTGGTATTCAGCAACGGG 3'.

The E3R primer (reverse) comprises

5' CGCGCGGGTTCTCGTCCTGGGTGTTTTCCTTCTCCTGCCCCCGC  
ACTCGGGATGCG 3', wherein a preferred fragment  
thereof further comprises

5' TCGTCCTGGGTGTTTTCCTTC 3'.

Using these primers in a second amplification  
reaction step a segment of CYP2A6 exon 3 is specifically  
amplified. The method further comprises use of the  
restriction endonuclease *XcmI* to detect the CYP2A6v1  
mutation and *DdeI* to detect the CYP2A6v2 mutation.

According to a yet further aspect of the  
invention there is provided a kit for performing the afore  
described methods which kit includes at least a portion of  
DNA in accordance with the invention and preferably at  
least one control sample of DNA containing the mutation or

- 25 -

° mutations of interest and ideally also a wild-type sample of DNA so that suitable comparisons can be made.

It is of note that although the method is described with reference to the above methods, any suitable method using the genetic material of the invention may be used to identify the mutations described herein.

The CYP2C9 assay has been used in a study of warfarin dose requirement in 94 patients undergoing anticoagulant treatment and the results obtained are summarized in Figure 5. 58 patients (61.7%) were homozygous for the wild-type (Arg<sub>144</sub>) allele and were found to require a median weekly maintenance dose of 31.5 mg of warfarin. 36 patients (38.6%) were heterozygous and required a median weekly maintenance dose of 24.5 mg. The doses required by the two groups were significantly different (Mann-Whitney U-test,  $p = 0.016$ ). No subjects in the group were homozygous for the mutant allele but based on allele frequencies and the Hardy Weinberg equilibrium, the predicted frequency of homozygous mutant subjects is 3.7%.

Comparison of the weekly maintenance dose of warfarin in the R144C heterozygotes ( $n = 36$ ) and homozygous wild-type ( $n = 58$ ) reveals that the heterozygotes required a significantly lower dose (range of 10.5 - 80.mg). Moreover, of the patients requiring the lowest doses to maintain an anticoagulation target (INR 2.0-4.0), in the range 5-15 mg per week, 9 out of 10 were heterozygous. At the other extreme of weekly doses >55 mg, 5 out of 6 patients were homozygous wild-type for CYP2C9. The significantly lower (20%) warfarin dose requirement of the patients with one variant R144C allele is consistent with the kinetic properties of the R144C protein with respect to (S)-warfarin hydroxylation and presumed in vivo metabolic clearance (Rettie et al. 1994 *Pharmacogen.*, 4:39-42).

The CYP2A6 genotyping assay has been used in

- 26 -

studies on coumarin metabolism. Coumarin 7-hydroxylase activity is a convenient marker activity to identify the presence of CYP2A6 in a particular sample. There is considerable variation in the ability of individuals to 7-hydroxylate this compound which is a reaction specific for CYP2A6. A subject deficient in coumarin 7-hydroxylation has been identified. This subject is homozygous for the mutant CYP2A6v1 allele confirming the previous *in vitro* findings that substitution of Leu160 by His results in loss of coumarin 7-hydroxylase activity. As shown in Fig. 6, CYP2A6 genotyping and phenotyping with coumarin has been performed on other members of the proband's family and impaired coumarin 7-hydroxylation has been observed in heterozygotes for the CYP2A6v1 mutation.

The genotyping assays described herein resulted from a two step amplification reaction wherein first amplification reaction amplifies a 7.8 Kb fragment containing the CYP2A6 gene (Fig. 9A) and a second amplification reaction amplifies an exon-specific fragment of CYP2A6. The amplification product was digested with restriction endonucleases producing different patterns for the various CYP2A6 alleles. Representative results obtained for several human subjects for the detection of the CYP2A6v1 (*XcmI* digestion) and CYP2A6v2 (*DdeI* digestion) are shown in Figure 9 panel B. A schematic depiction of this genotyping assay is shown in Figure 9, panel C. Of 155 human genomic DNA samples analyzed 21 heterozygous (+/-) and 6 homozygous (-/-) subjects were detected for the CYP2A6v1 allele, whereas 17 heterozygous (+/-) and no homozygous were identified for the CYP2A6v2 allele variant. Additionally, 7 homozygous for both CYP2A6v1 and CYP2A6v2 alleles were found.

Allelic frequencies were calculated for either allele in several ethnic groups and analyzed as shown in Table 1. CYP2A6v1 frequency is almost identical between Caucasian and Japanese, and it is only twice the frequency in Taiwanese samples. Significantly, this allele is

- 27 -

completely absent in the African-American population within the samples studied. The Japanese population has a remarkable higher frequency for the CYP2A6v2 allele (28%) as compared to the Caucasian (2%), Taiwanese (6%) or African-American (2.5%) (ethnic groups).

Table 1: Allelic frequency for the CYP2A6 gene in different ethnic groups.

Ethnic Group	Allelic Frequencies (%)			
	CYP2A6	CYP2A6v1	CYP2A6v2	N
Caucasian	75	23	2	52
Japanese	52	20	28	40
Taiwanese	83	11	6	178
African-American	97.5	0	2.5	40

The following examples illustrate various aspects of the present invention and in no way are intended to limit the scope thereof. All books, articles, and patents referenced herein are incorporated herein, in toto, by reference. Other similar embodiments will be clear to the skilled artisan and are encompassed within the spirit and purview of the present invention.

#### EXAMPLE 1

##### Method for determining the genotype CYP2C9

Genotyping for the CYP2C9 polymorphism is carried out by amplification by PCR followed by digestion with the restriction endonuclease *AvaII*. Amplifications are performed in 0.5 ml microcentrifuge tubes in a volume of 100  $\mu$ l containing 10 mM Tris-HCl, pH 8.8, 1.5 mM MgCl<sub>2</sub>, 50 mM KCl, 0.1% Triton X-100, 5% dimethylsulphoxide, 200  $\mu$ M each of dTTP, dATP, dCTP and dGTP, 250  $\mu$ M of the primers HF18 and HF2R, 2.5 units Taq polymerase and 1  $\mu$ g human leukocyte genomic DNA. PCR conditions consist of 35 cycles with a denaturation at 93°C for 1 min. annealing at 55°C for 1.5 min and polymerization at 72°C for 1 min. 20  $\mu$ l of the amplified DNA is incubated with 10 units *AvaII* for 3h at 37°C and then analyzed by electrophoresis on

- 28 -

1.8% agarose minigels in TBE (90 mM Tris-borate, 2 mM EDTA) buffer. The digestion products are visualized by ethidium bromide staining. DNA from individuals positive for the wild-type Arg<sub>144</sub> is digested to give fragments of 270 bp and 50 bp whereas in individuals with the mutant Cys<sub>144</sub> present, a band of 320 bp is seen due to loss of an AvaII site (Figure 3).

#### EXAMPLE 2

Genotyping for the CYP2C9 polymorphism was carried out by amplification by PCR followed by digestion with the restriction endonuclease AvaII.

One hundred patients were recruited from two anticoagulation clinics in the Newcastle area over four study days. Body weight and height were measured, the basal metabolic index ("BMI") calculated for each patient and details of age, sex, drug history, current and previous International Normalized Ratio ("INR") determinations, indications for anticoagulation and other significant health problems were all recorded. DNA was isolated by a standard manual chloroform-phenol extraction procedure and 1µg was subjected to PCR analysis. As shown in Figure 10 the C→T substitution, which converts Arg<sub>144</sub> to Cys, resides in exon 3 of the CYP2C9 gene and results in the loss of an AvaII restriction site (...GAGGACCGTGTTCAA...) in the R144C allele (...GAGGACTGTGTTCAA...). This provided the basis of the amplification strategy. A CYP2C9 specific intron forward primer (HF18, TGCAAGTGCCTGTTTCAGCA, Figure 10) and a CYP2C9 exon 3 3'-end reverse primer (HF2R, AGCCTTGTTTTTCTCAACTC, Figure 10) were used at a concentration of 250µM each. Amplifications were performed in a volume of 100 µl containing 20 mM Tris HCl (pH 8.3), 1.5 mM MgCl<sub>2</sub>, 25 mM KCl, 0.05% (w/v) Tween 20, 10 µg gelatin/ml, 2% (w/v) DMSO, 200 µM each of dATP, dCTP, dGTP and dTTP and 2.5 units of Taq DNA polymerase (Perkin-Elmer). Reactions were carried out for 35 cycles

- 29 -

at an annealing temperature of 55°C for 90 sec, a polymerase temperature of 72°C for 1 min, and a heat denaturing temperature of 93°C for 1 min, using a Perkin-Elmer Cetus DNA thermal cycler. The PCR products digested with AvaII and sized using NuSieve agarose gels (3% NuSieve, 0.75% agarose). Presence of the CYP2C9 wild-type and R144C alleles were detected as fragments of 50 + 270 bp and 320 bp respectively (see Figures 3). The PCR product synthesized from human genomic DNA with the primers HF18/HF2R was directly sequenced on an ABI 373A automatic sequencer. Briefly, the PCR product was first purified by using the Wizard DNA clean-up system (Promega Co., Madison, WI). The purified template was then subjected to dideoxy terminator cycle-sequencing with the primers HF18 and HF2R. The primer-extended products were purified and sequenced following the manufacturer's procedure. Sequence analysis was done by using the MacVector software program (Eastman-Kodak Co., Rochester, NY).

DNA was obtained from 94 patients. Of these 58 (62%) were homozygous for the wild-type CYP2C9 gene and 36 (38%) were heterozygous for the R144C allele. No R144C homozygotes were found. The frequency of the wild-type (Arg-144) and R144C (Cys-144) alleles in the study population is thus 0.808 and 0.192 respectively. An expectation of 3.7% R144C homozygotes can be anticipated from the Hardy-Weinberg equilibrium, but the 95% confidence interval in this estimation of 0.8-8.4% and thus the finding of zero homozygotes in 94 patients is not significantly different from expectation. The specificity of the PCR reaction with respect to the CYP2C9 gene was confirmed by sequencing. The alignment of the sequence obtained from the PCR product with that corresponding to the CYP2C9 gene showed a 100% degree of homology. Interestingly, a heterozygous pattern was obtained for the R144C allelic variant, confirming the high frequency of this allele within the normal population. No sequence

- 30 -

deriving from CYP2C9, CYP2C18 or CYP2C19 was found confirming the specificity of the assay for CYP2C9.

### EXAMPLE 3

#### Method for determining the genotype CYP2A6

Genotyping for the CYP2A6 polymorphism is carried out by allele-specific PCR using two parallel PCR reactions, one specific for the wild-type allele, one for the mutant allele. Amplifications are performed in 0.5 ml microcentrifuge tubes in a volume of 45  $\mu$ l containing 10 mM Tris-HCl, pH 8.8, 1.5 mM MgCl<sub>2</sub>, 50 mM KCl, 0.1 % Triton X-100, 5 % dimethylsulfoxide, 200  $\mu$ M each of dTTP, dATP, dCTP and dGTP, 250  $\mu$ M of the primers B and either J51 or J61, 1.25 units Taq polymerase and 1  $\mu$ g human leukocyte genomic DNA. PCR conditions consist of 40 cycles with a denaturation at 93°C for 1 min., annealing at 57°C for 2 min and polymerization at 70°C for 2 min. The products are analyzed by electrophoresis on 1% agarose minigels in TBE buffer and DNA is visualized by staining with ethidium bromide. As shown in Figure 4, there are three possible results: the individual may be homozygous for the wild-type allele and give a DNA product only for the PCR reaction with primer J51, the individual may be heterozygous with one wild-type and one mutant allele and give DNA products with both primers J51 and J61 or the individual may be homozygous for the mutation and give a DNA product only with the J61 primer.

### EXAMPLE 4

#### Alternative Method for Determining the Genotype CYP2A6

For use of F4 and R4 primers, each reaction mixture contained 600 ng human genomic DNA, 0.2  $\mu$ M of each primer, 200  $\mu$ M dNTP's, 0.8 mM magnesium acetate and 2 units of rTth I DNA polymerase. Hot start was as indicated by the manufacturer (Perkin Elmer) and the amplification reaction of 31 cycles of 93°C, 1 min; 66°C, 6 min 30 sec. Amplification products were analyzed in

- 31 -

0.7% agarose gels and the DNA visualized by staining with ethidium bromide. For the exon 3 specific amplification, the reaction which uses, the primers E3F and E3R consist of 5 $\mu$ l of the 7.8 Kb PCR reaction, 0.5  $\mu$ M of each primer, 200  $\mu$ M dNTP's, 1.5 mM MgCl<sub>2</sub> and 2.5 units of Taq DNA polymerase. The amplification reaction consisted of 94°C for 3 minutes followed by 31 cycles of 94°C, 1 minute; 60°C, 1 minute and 72°C, 1 minute.

Amplification products were then digested without purification with restriction endonucleases which detect the CYP2A6 wild type (no digestion), CYP2A6v1 (XcmI) and CYP2A6v2 (DdeI). DNA was visualized by use of ethidium bromide after electrophoresis in 1% agarose, 3% NuSieve agarose.

It is of note that CYP2C9 genotyping can be performed using an allele-specific assay similar to that used above for CYP2A6.

- 32 -

CLAIMS

1. A CYP2A6v2 DNA having a coding sequence shown in Figure 11.

5 2. The DNA of claim 1 having a genomic sequence as shown in Figure 12.

10 3. A CYP2A13 DNA having a coding sequence shown in Figure 13.

4. The DNA of claim 3 having a genomic sequence shown in Figure 14.

15 5. A nucleic acid primer sequence comprising at least ten (10) contiguous nucleotide bases selected from the sequence showing in Figure 12.

20 6. A nucleic acid primer sequence comprising at least ten (10) contiguous nucleotide bases selected from the sequence shown in Figure 14.

7. A nucleic acid primer sequence selected from the group consisting of:

- 25 A. 5' GGCTTCCTCATCGACGCACT 3';  
B. 5' GGCTTCCTCATCGACGCACA 3';  
C. 5' AATTCCAGGAGGCAGGGCCT 3';  
D. 5' TGCAAGTGCCTGTTTCAGCA 3';  
E. 5' AGCCTTGGTTTTTCTCAACTC 3';  
F. 5' CCCCTTATCCTCCCTTGCTGGCTGTGTCCCAAGCTAGGCA  
30 GGATTCATGGTGGGGCA 3';  
G. 5' GCCACCACGCCCCTTCCTTTCCGCCATCCTGCCCCCAGTC  
TTAGCTGCGCCCCTCTC 3';  
H. 5' CCTGATCGACTAGGCGTGGTATTCAGCAACGGGGAGCGCG  
CCAAGCAGCTCCTG 3';  
35 I. 5' CGCGCGGGTTCCTCGTCCCTGGGTGTTTTCTTCTCCTGCC  
CCCGCACTCGGGATGCG 3';

- 33 -

°  
or any nucleic acid sequence of at least 10 contiguous nucleotides selected from any one of A-I.

5 8. A method of determining the presence or absence of an allelic variant in CYP2A6 or CYP2C9 DNA comprising:

- 10 (a) amplifying an exon containing a variant sequence with in said DNA, producing an extension product;
- (b) treating extension products with at least one restriction endonuclease under conditions sufficient to produce digestion fragments;
- 15 (c) analyzing the digestion fragments, for a variant specific digestion fragment or lack thereof.

9. The method of claim 8 wherein a CYP2C9 variant DNA is being detected.

20 10. The method of claim 9 wherein the amplifying step is a polymerase chain reaction using primers comprising HF18 and HF2R.

25 11. The method of claim 8 wherein step (a) is preceded by a gene-specific amplification reaction.

12. The method of claim 11 wherein the gene-specific amplification is a polymerase chain reaction.

30 13. The method of claim 12 wherein a CYP2A6 variant is being detected.

35 14. The method of claim 13 wherein a gene-specific amplification reaction uses primers comprising F4 and R4 and the exon amplification reaction uses primers comprising E3F and E3R.

- 34 -

15. The method according to claim 10 wherein the extension products are treated with the restriction endonuclease *AvaII*.

5 16. The method according to claim 14 wherein the extension products are treated with at least one restriction endonuclease comprising *DdeI* and *XcmI*.

10 17. A method of determining the presence or absence of an allelic variant in CYP2A6 or CYP2C9 DNA comprising:

- 15 (a) contacting said DNA with a first primer encompassing a nucleotide variation specific to variant DNA and a second primer which is complementary to a region of said DNA such that upon hybridization and amplification, an extension product will be formed;
- 20 (b) analyzing the extension products for allelic-variant specific extension products.

18. The method of claim 17 wherein a CYP2A6 variant DNA is being detected.

25 19. The method of claim 18 wherein the amplifying step is a polymerase chain reaction wherein the first primer comprises J51 and J61 and the second primer comprises primer B.

30 20. A kit for determining the presence or absence of an allelic variant of CYP2A6 or CYP2C9 DNA comprising: at least one nucleic acid primer sequence capable of hybridizing to said DNA; the kit further containing instructions relating to the determination of  
35 the presence or absence of an allelic variant of CYP2A6 or CYP2C9 DNA.

- 35 -

21. The kit according to claim 20 further comprising amplification components and at least one restriction endonuclease.

22. The kit of claim 20 wherein the CYP2A6 allelic variant is being detected.

23. The kit of claim 22 wherein the nucleic acid primers comprise F4, R4, E3F and E3R.

24. The kit according to claim 20 wherein the CYP2C9 allelic variant is being detected.

25. The kit according to claim 25 wherein the nucleic acid primers comprise HF18 and HF2R.

26. A process for providing a human with a therapeutic CYP2A6v2 or CYP2A13 DNA segment said human cells expressing in vivo in said human or therapeutically effective amount of said protein.

27. A pharmaceutical composition comprising an antisense nucleic acid derived from CYP2A6v2 DNA.

28. A pharmaceutical composition comprising and antisense nucleic acid derived from CYP2A13.

0 1 / 2 9

.....exon 2.....

IIc1 (C9)	GATCTTGGAGAGGAGTTTTCTGGAAGAGGCCATTTTCCCACTGGCT
	Asp Leu Gly Glu Glu Phe Ser Gly Arg Gly Ile Phe Pro Leu Ala
IIc2 (C8)	GATAATGGAGAGGAGTTTTCTGGAAGAGGCCAATTTCCCAATATCT
	Asp Asn Gly Glu Glu Phe Ser Gly Arg Gly Asn Ser Pro Ile Ser
DIIC2	
Clone 4 (hIIc1-4)	GATCTTGGAGAGGAGTTTTCTGGAAGAGGCCATTTCCCACTGGCTG
Clone 18 (hIIc1-18)	GATCATGGAGAGGAGTTTTCTGGAAGAGGAAGTTTCCAGTGGCTG
Clone 3	
Clone 16	GATCATGGAGAGGAGTTTTCTGGAAGGAAGTATTTCCAGTATCCA
Clone 21 (hIIc1-21)	GATCTTGGAGAGGAGTTTTCTGGAAGAGGCCATTTCCCACTGGCTG
Clone 26 (hIIc1-26)	GATCATGGAGAGGAGTTTTCTGGAAGAGGAAGTTTCCAGTGGCTG
Clone 33	

**FIG. 1** (Sheet 1)

0 2 / 2 9

IIIC1	.....exon 2.....][Start of intron 2.....
	GAAAGAGCTAACAGAGGATTTG
	Glu Arg Ala Asn Arg Gly Phe G
IIIC2	CAAAGAATTACTAAAGGACTTG
	Gln Arg Ile Thr Lys Gly Leu G
DIIC2	CTTGTAAGTGCACATATTCTGTGTACGCTTGGTAC
Clone 4	AAAGAGCTAACAGAGGATTGGTAGGTGIGCAAGIGCCTGHHICAGCAICTGTCTTGG
	Primer HF-18
Clone 18	AAAAAGTTAACAAAGGACTTGTAATGTGCATGTATCGTGTGTATGTGTACATGT
Clone 16	AAAAAAGCTA GTAAGGAGTTGGTACATGTGTGTGTCAGTGTGTGCGCCTTGTCTG
Clone 21	AAAGAGCTAACAGAGGATTGGTAGGTGTGCAAGTGCCCTGTTTCAGCAICTGTCTTGG
Clone 26	AAAAAAGTTAACAAAGGACTTGTAATGTGCATGTATCGTGTGTATGTGTACATGT

FIG. 1 (Sheet 2)

.....Intron 2.....

0 3 / 2 9

DIIC2	TCGGGTGAGGGGGATGGAACACAGAGCCCTAAAAAGCTTCTCAGCAGAGCTTAGC
Clone 4	GGATGGGGAGGATGGAACACAGAGACTTACAGAGCTCCTCGGGCAGAGCTTGGCCCA
Clone 18	GTATGTACTGGGCAGTGGCTATAGGGATGGGGAGGATGGAACACAGGCTTGAAAA
Clone 3	CAGAAGGTGAAT(G)GAAACAACACAC(T)TGAA
Clone 16	TATTAGTAATGAGGCAGAGGTGAATGGAACAAACACTTGAAGAGCTCCTAAA
Clone 21	GGATGGGGAGGATGGAACACAGAGCTA CTA GCAGAGCT(T)]CTCGGG
Clone 26	GTATGTACTGGGCAGTGGCTATAGGGATGGGGAGGATGGAACACAGGCTTGAAGA

**FIG. 1** (Sheet 3)

.....intron 2.....

DIIC2      CTATCTGCATGGCTGCCAAGTGTGACGACATTTCTTCCTTGGCTGTGAATTCTC  
Clone 4    TCCACATGGCTGCCCCAGTGTACAGCTTCCTCTTTCTTGGCTGGGATCTCCCTCCTA  
Clone 18   GCTCCTGGGACAGAACTTGACCTGTCCACGTGGCTGCCGAGTGTACGCTCTCTTG  
Clone 3    GAGCTCCTAAAC(T)TAGC(T)TAGCTTGGCCATTGGGCTGTGAAATCAGCTTC  
Clone 16   ACTTAGCTTGGCC(C)ATTTGGTGGCTGTGAAATCAGCTTCCCTTCNNNC(C)TGG  
Clone 21   CAGAGCTTGGCCCATCCACATGGCTGCCCCAGTGTACGCTTCTTCTTGCCCTG  
Clone 26   GCTCCTGGGACAGAACTTGACCTGTCCACGTGGCTGCCGAGTGTACGCTCTCTTG

0 4 / 2 9

**FIG. 1** (Sheet 4)

0 5 / 2 9

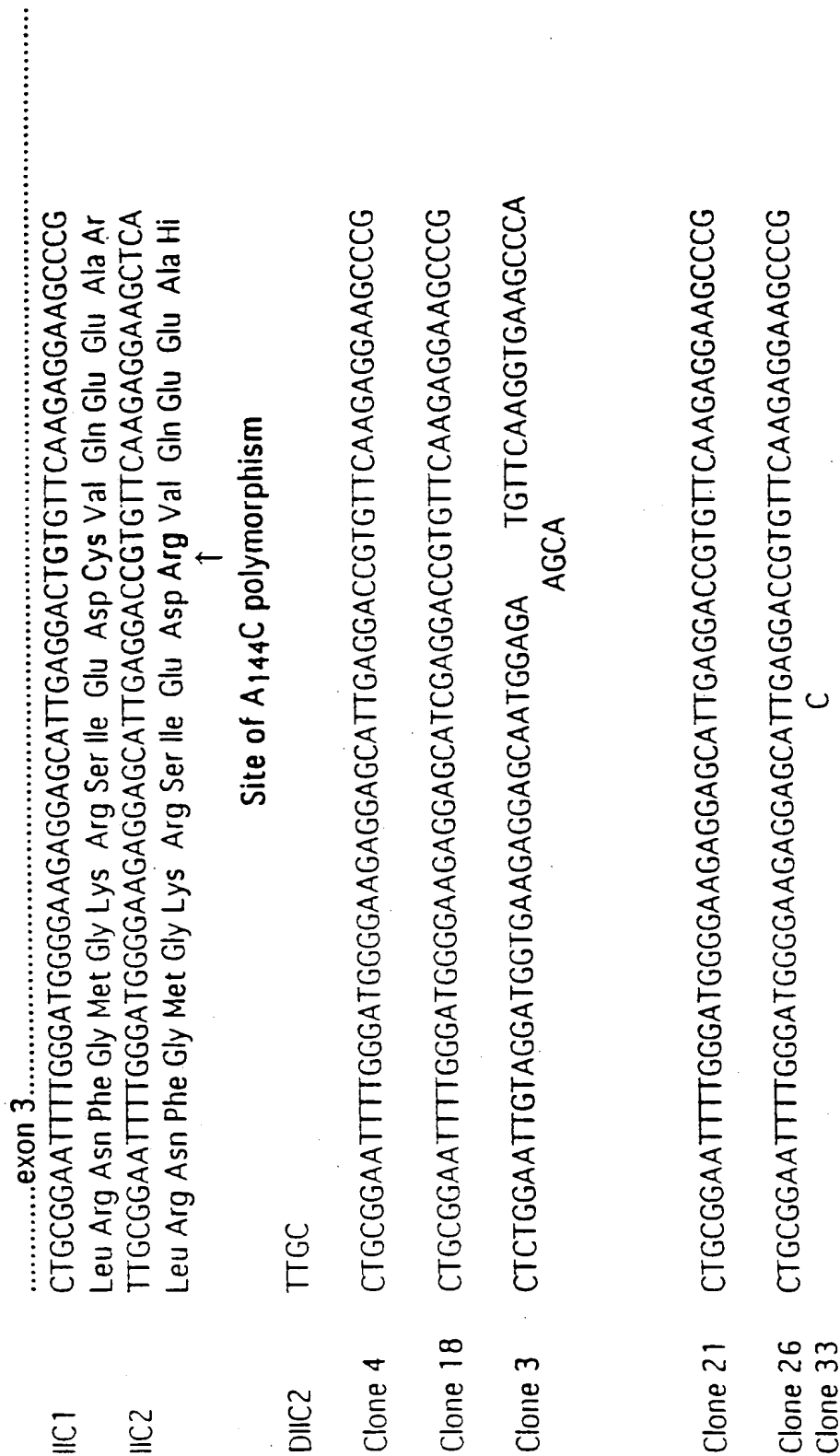
.....end of intron 2]

DIIC2	CCAGTTTCTGCCCTTTTTTATTAG	
Clone 4	GTTTCGTTTCTCTTCCCTGTTAG	
Clone 18	TCCTTGTTTGGATTCTCCCTCGTAGCTTCTGTTTTCTGTTCTGCTAG	
Clone 3	CTCTTTCTTGCCTGGGATCTCCCTCCCTCGTTTCTGTTTCCCTTCTTCA	
Clone 16	ATCTCCCTCCCTCGTTTCTGTTCCCTTCTC	A
Clone 21	GGATCTCCCTCCTAGTTTTCGTTTCTCTTCCCTGTT	AG
Clone 26	TCCTTGTTTGGATTCTCCCTCGTAGCTTCTGTTTTCTGTTCTGCTAG	

**FIG. 1** (Sheet 5)

[Start of exon 3.....	
IIC1	GAATTGTTTTCAGCAATGGAAAGAAATGGAAGGAGATCCGGCGTTTCTCCCTCATGACG ly Ile Val Phe Ser Asn Gly Lys Lys Trp Lys Glu Ile Arg Arg Phe Ser Leu Met Thr
IIC2	GAATCATTTTCCAGCAATGGAAAGAGAGATGGAAAGGAGATCCGGCGTTTCTCCCTCACAACC ly Ile Ile Ser Ser Asn Gly Lys Arg Trp Lys Glu Ile Arg Arg Phe Ser Leu Thr Thr
DIIC2	GAATCATTTTCCAGCAATGGAAAGAGATGGAAAGGAGATCCGGCGTTTCTCCCTCACAACC
Clone 4	GAATTGTTTTCAGCAATGGAAAGAAATGGAAGGAGATCAGGCGTTTCTCCCTCATGACG
Clone 18	GAATCCTTTTTCAGCAATGGAAAGAGAGATGGAAAGGAGATCCGGCGTTTCTGCCTCATGACT
Clone 3	GGATCATTTTTTAGCAATGGAAAGAGAGATGTAAGGATGCTGGCTCTTCTTGCTCATGACG
Clone 16	GGATCATTTT
Clone 21	GAATCGTTTTCAGCAATGGAAAGAGAGATGGAAAGGAGATCCGGCGTTTCTCCCTCATGACG
Clone 26	GAATCCTTTTTCAGCAATGGAAAGAGAGATGGAAAGGAGATCCGGCGTTTCTCCCCCATGACG
Clone 33	G T T

FIG. 1 (Sheet 6)



**FIG. 1** (Sheet 7)

08/29

2A6 Intron 2	-----GT G C C C A A G A G G G G A A G G T G G C C A G G T G G A C T C A A G G T C T	42
2A8 Intron 2	GT G A G G G G G T G C C C A A G A G G G G A A G G T G G C C A G G T G G A C T C A A G G T C T	49
2A7 Intron 2	-----GT G C C C A A G A G G G G A A G G T G G C C A G G T G G A C T C A A G G T C T	41
Consensus	GT G A G G G G G T G C C C A A G A G G G G A A G G T G G C C A G G T G G A C T C A A G G T C T	50
2A6 Intron 2	C G T G T T G C C A G C C T T C T C C C C T G A C T C T C C T G C C A C T G G A G G T T A T G G G	92
2A8 Intron 2	C G T G T T G C C A G C C T T C T C C C C T G A C T C T C C T G C C A C T G G A G G T T A T G G G	99
2A7 Intron 2	C G T G T T G C C A G C C T T C T C C C C T G A C T C T C C T G C C A C T G G A G G T T A T G G A	91
Consensus	C G T G T T G C C A G C C T T C T C C C C T G A C T C T C C T G C C A C T G G A G G T T A T G G	100
2A6 Intron 2	A G A G T C C C C C A G C C T G G T C T C C T C C C C A T C T C C C T C A T T G G G G C T C T	142
2A8 Intron 2	A G A G T C C C C C A G C C T G G T C T C C T C C C C A T C T C C C T C A T T G G G G C T C T	149
2A7 Intron 2	A G A G T C C C C C A G C C T G G T C T C C T C C C C A T C T C C C T C A T T G G G G C T C T	141
Consensus	A G A G T C C C C C A G C C T G G T C T C C T C C C C A T C T C C C T C A T T G G G G C T C T	150
2A6 Intron 2	C C T G T G T A T C C C C C A C C T G T C T C C A G C C C C C C G T G A T T C C T C C C T	192
2A8 Intron 2	C C T G T G T A T C C C C C A C C T G T C T C C A G C C C C C C G T G A T T C C T C C C T	199
2A7 Intron 2	C C T G T G T A T C C C C C A C C T G T C T C C A G C C C C C C G T G A T T C C T C C C T	190
Consensus	C C T G T G T A T C C C C C A C C T G T C T C C A G C C C C C C G T G A T T C C T C C C T	200
2A6 Intron 2	C C C T C T C T C T G C C C C A C C T C G A T T T C T C T C T C A T T G G A G T C T C T C T T	242
2A8 Intron 2	C C C T C T C T C T G C C C C A C C T C T C G A T T T C T C T C T C A T T G G A G T C T C T C T A	249
2A7 Intron 2	C C C T C T C T C T G C C C C A C C T C C A G A T T T C T C T C A T T G G A G T C T C T C T T	239
Consensus	C C C T C T C T C T G C C C C A C C T C G A T T T C T C T C T C A T T G G A G T C T C T C T T	250
2A6 Intron 2	C C C T C T C T C T T C C A T T C T C A A G G A C A T C T G G G T T T C T C T T T A C C A G C C C	292
2A8 Intron 2	C C C T C T C T C T T C C A T T C T C A A G G A C A T C T G G G T T T C T C T T T A C C A G C C C	299
2A7 Intron 2	C C C T C T C T C T T C C A T T C T C A A G G A C A T C T G G G T T T C T C T T T A C C A G C C C	267
Consensus	C C C T C T C T C T T C C A T T C T C A A G G A C A T C T G G G T T T C T C T T T A C C A G C C C	300
2A6 Intron 2	T G G G T C T C T G T C T A C A T G A G T C T T T G A G C C C T C T T A G C T T C T G G G T T C	342
2A8 Intron 2	T G G T C C T C T G T C T T C A T T T G T C T T T T T G C G C T C T G G G C T C T G T G C T T C	349
2A7 Intron 2	-----T G G T C C T C T G T C T T C A T T T G T C T T T T T G C G C T C T G G G C T C T G T G C T T C	271
Consensus	T G G K Y C T C T G T C T W C A T K W G T C T T T K A G K C S C T C T Y A G C T C T G K G C T T C	350
2A6 Intron 2	T C T G G G T T T C T C A T C T C T C C G G A T C C C T T T C T C A A T T T C T C C T C T C T T	392
2A8 Intron 2	T C C T T G T T T C T C C T C T C T C T G G T T C C C T C T C C A C T T C T C C T C T C T T	399
2A7 Intron 2	-----T C C T T G T T T C T C C T C T C T C T G G T T C C C T C T C C A C T T C T C C T C T C T T	271
Consensus	T C Y G K G T T T C T C T C T C T C Y G S W T C C C T Y T C Y C A M T T C T C C T C T C T T	400
2A6 Intron 2	A G G A T G C C A G G G T A T T C C T A C T T C C A C A T C T T C A G G C T C C A T C T C C T G G	442
2A8 Intron 2	A G G A T T T C A G G G T A T T C C T A C T T C C A C A T C T T C A G G C T C C C A A C T C C T G G	448
2A7 Intron 2	-----A G G A T T T C A G G G T A T T C C T A C T T C C A C A T C T T C A G G C T C C C A A C T C C T G G	271
Consensus	A G G A T K Y C A G G G T A T T C C T A C T T C C A C A T C T Y C A G G T Y C C A W C T C C T G G	450
2A6 Intron 2	T A A C A G T C T C T C T T C C C T C C A G C C C T C T G T T T T C T A T C A A T A T T A A	492
2A8 Intron 2	T A A T T G T C T C T C T T C C C T C C C T C T C G A T C C C T C T G T T T T C T A T T C A T A T T T	498
2A7 Intron 2	-----T A A T T G T C T C T C T T C C C T C C C T C T C G A T C C C T C T G T T T T C T A T T A T T T	271
Consensus	T A A T W G T C T C T C Y T C C T T C C M G A Y C C T C T C T G T T T T C T A T A T A T T W	500
2A6 Intron 2	A C T C T C T T T G C T C A G G T C A G G T T A A G A A T C T C A C A C C A A G A G A G S A T G T	540
2A8 Intron 2	T C T C T C T C T T C T C C A G T T C A G A T T A A G A A T C T T C A C C A A T T T T T A T T T	546
2A7 Intron 2	-----A C T C T C T C T T C T C C A G T T C A G A T T A A G A A T C T T C A C C A A T T T T T A T T T	271
Consensus	W C T C T C T C T K C T C C A G T T C A G A T T A A G A A T C T Y W C A C C A A G A K W K K A T K T	550

FIG. 2 (Sheet 1)

0 9 / 2 9

2A6 intron 2	CCTCCACCCA	GATCTCCCCA	TATCTCACTA	CCCCACCCCTC	CATC---CTC	587
2A8 intron 2	CCTCCTCCCA	GATCTCCCCA	TATCTCACTT	CCCCTCCCTC	CATCTCTCTC	596
2A7 intron 2	-----	-----	-----	-----	-----	271
Consensus	CCTCCWCCCA	GATCTCCCCA	TATCTCACTW	CCCCWCCCTC	CATCTCTCTC	600
2A6 intron 2	TGCCT----C	CATCAC--TC	TCTTTCTC--	-----TCC	CC--A-----	613
2A8 intron 2	TTTCTCTCCC	CACTACCTTC	CCTTCCTCCA	TGGAGTATCC	CCGTATCCCT	646
2A7 intron 2	-----	-----	-----	-----	-----	271
Consensus	TKYCTCTCCC	CAYYACCTTC	YCTTYCTCCA	TGGAGTATCC	CCGTATCCCT	650
2A6 intron 2	CTGCCCCCTGC	GGACGCGATC	CAATGG--AG	TGTG-----	----GA---G	650
2A8 intron 2	CTGTTTCTCT	GCATCTGTCT	GTCTGGCCTT	TCTGCTTCTC	TTCTGATTCT	696
2A7 intron 2	-----	-----	-----	-----	-----	271
Consensus	CTGYYYCTSY	GSAYSYGWYY	SWMTGGCCWK	TSTGCTTCTC	TTCTGATTCK	700
2A6 intron 2	CTAATGCCGT	-----GAA	GCTATGTGCA	TCTCTCTGTC	TGGCCGTACC	693
2A8 intron 2	CTTATTCTTT	CTACCCGGAC	TCTCTCTCTC	TCTCTCTCTC	TCTCTCTCTC	746
2A7 intron 2	-----	-----	-----	-----	-----	271
Consensus	CTWATKCYKT	CTACCCGGAM	KCTMTSTSYM	TCTCTCTSTC	TSKCYSTMYC	750
2A6 intron 2	TGGGT---AA	TAACCTGATC	GACT-----	-----	-----	714
2A8 intron 2	TCTCTCTCTC	TCTCTCTCTC	TCTCTCTCTC	TCTCTCTCTA	TATATATATA	796
2A7 intron 2	-----	-----	-----	-----	-----	271
Consensus	TSKSTCTCWM	TMWCYYKMTC	KMYTCTCTC	TCTCTCTCTA	TATATATATA	800
2A6 intron 2	-----	-----	-----	-----	-----	714
2A8 intron 2	TATATATATA	CACACACACA	CACACACACA	CACACACACA	CACACACATA	846
2A7 intron 2	-----	-----	-----	-----	-----	271
Consensus	TATATATATA	CACACACACA	CACACACACA	CACACACACA	CACACACATA	850
2A6 intron 2	-----	-----	-----	-----	-----	714
2A8 intron 2	TATATTAGGG	GGGGACTCCC	TTTCTGCTCC	ACCCTTGGGG	AGCCCCTTGG	896
2A7 intron 2	-----	-----	-----	-----	-----	271
Consensus	TATATTAGGG	GGGGACTCCC	TTTCTGCTCC	ACCCTTGGGG	AGCCCCTTGG	900
2A6 intron 2	-----	-----	-----	-----	-----	714
2A8 intron 2	AACTGGTCCG	CTCTGCTACC	ACCACCCCCT	GACCTCTCTC	CACCCCCGCG	946
2A7 intron 2	-----	-----	-----	-----	-----	271
Consensus	AACTGGTCCG	CTCTGCTACC	ACCACCCCCT	GACCTCTCTC	CACCCCCGCG	950
2A6 intron 2	-----	---	---	---	---	714
2A8 intron 2	TTCACCTCCC	CA				958
2A7 intron 2	-----	---	---	---	---	271
Consensus	TTCACCTCCC	CA				962

Intron 2 alignment

FIG. 2 (Sheet 2)

10 / 29

2A8 exon 3	GCCTGGCCTT CAGCAACGGG GAGCGCGCCA AGCAGCTCCG GCGCTTCACG	50
2A6 exon 3	GCCTGGCTTT CAGCAACGGG GAGCGCGCCA AGCAGCTCCG GCGCTTCACG	50
2A7 exon 3	GCCTGGCCTT CAGCAACGGG GAGCGCGCCA AGCAGCTCCG GCGCTTCACG	50
Consensus	GCCTGGCTTT CAGCAACGGG GAGCGCGCCA AGCAGCTCCG GCGCTTCACG	50
2A8 exon 3	ATCGCCACCC TTAGGGGTTT TGGGTGGGGC AAGCGGGCA TCGAGGATCG	100
2A6 exon 3	ATCGCCACCC TTAGGGGTTT TGGGTGGGGC AAGCGGGCA TCGAGGATCG	100
2A7 exon 3	ATCGCCACCC TTAGGGGTTT TGGGTGGGGC AAGCGGGCA TCGAGGATCG	100
Consensus	ATCGCCACCC TTAGGGGTTT TGGGTGGGGC AAGCGGGCA TCGAGGATCG	100
Codon 160		
2A8 exon 3	CATCCAGGAG GAGTCGGGCT TCCTCATCGA GGCCATCCGG AGCAGGCACG	150
2A6 exon 3	CATCCAGGAG GAGTCGGGCT TCCTCATCGA GGCCATCCGG AGCAGGCACG	150
2A7 exon 3	CATCCAGGAG GAGTCGGGCT TCCTCATCGA GGCCATCCGG AGCAGGCACG	150
Consensus	CATCCAGGAG GAGTCGGGCT TCCTCATCGA GGCCATCCGG AGCAGGCACG	150
← Primer J51/61 →		

## Exon 3 alignment

FIG. 2 (Sheet 3)

11 / 29

2A8 intron 3	-----	-----	-----	G	AGAGTAGGG	ACCCGAGTG	21	
2A6 intron 3	GTGAGCAGGG	GACCCCGAGT	GCGGGGGCG	G	GAGHAGGA	ACCCGAGTG	44	
2A7 intron 3	GTGAGTAAGG	TTCCCGGAGT	GCGGGGGCG	G	GAGHAGGA	ACCCGAGTG	44	
Consensus	GTGAGYARGG	KWCCCGGAGT	GCGGGGGCG	G	AGAGTAGGG	ACCCGAGTG	50	
2A8 intron 3	CCAGGCGCG	G	GAACCGCG	CGC	GCGTTCTGCC	TGGGGATGGG	GACTAGGTGG	68
2A6 intron 3	CCAGGCGCG	G	GAACCGCG	CGC	GCGTTCTGCC	TGGGGATGGG	GACTAGGTGG	94
2A7 intron 3	CCAGGCGCG	G	GAACCGCG	CGC	GCGTTCTGCC	TGGGGATGGG	GACTAGGTGG	94
Consensus	CCAGGCGCG	G	GAACCGCG	CGC	GCGTTCTGCC	TGGGGATGGG	GACTAGGTGG	100
2A8 intron 3	GGAAAGGGGC	CCGCACCTTC	AGCCCTGGAG	TCTGGCGCT	GGGATTGGG	117		
2A6 intron 3	GGAAAGGGGC	CCGCACCTTC	AGCCCTGGAG	TCTGGCGCT	GGGATTGGG	144		
2A7 intron 3	GGAAAGGGGC	CCGCACCTTC	AGCCCTGGAG	TCTGGCGCT	GGGATTGGG	144		
Consensus	GGAAAGGGGC	CCGCACCTTC	AGCCCTGGAG	TCTGGCGCT	GGGATTGGG	150		
2A8 intron 3	TCAACAGGGC	CCTGCCTCCT	GGAATTCTGA	CTCTCCTCAG	ACCTCTGAGT	167		
2A6 intron 3	TCAACAGGGC	CCTGCCTCCT	GGAATTCTGA	CTCTCCTCAG	ACCTCTGAGT	194		
2A7 intron 3	TCAACAGGGC	CCTGCCTCCT	GGAATTCTGA	CTCTCCTCAG	ACCTCTGAGT	194		
Consensus	TCAACAGGGC	CCTGCCTCCT	GGAATTCTGA	CTCTCCTCAG	ACCTCTGAGT	200		
← Primer B →								
2A8 intron 3	TGACTCTCTC	CCCAACCCCC	CTTCTCCCGC	CATACCTGGA	-	207		
2A6 intron 3	TGACTCTCTC	CCCAACCCCC	T-TCTCCCGA	CATACCTGGA	-	233		
2A7 intron 3	TGACTCTCTC	CCCAACCCCC	TCTCTCCCGC	CATACCTGGA	G	235		
Consensus	TGACTCTCTC	CCCAACCCCC	YYTCTCCCGM	CATACCTGGA	G	241		

## Intron 3 alignment

FIG. 2 (Sheet 4)

.1.2 / 29

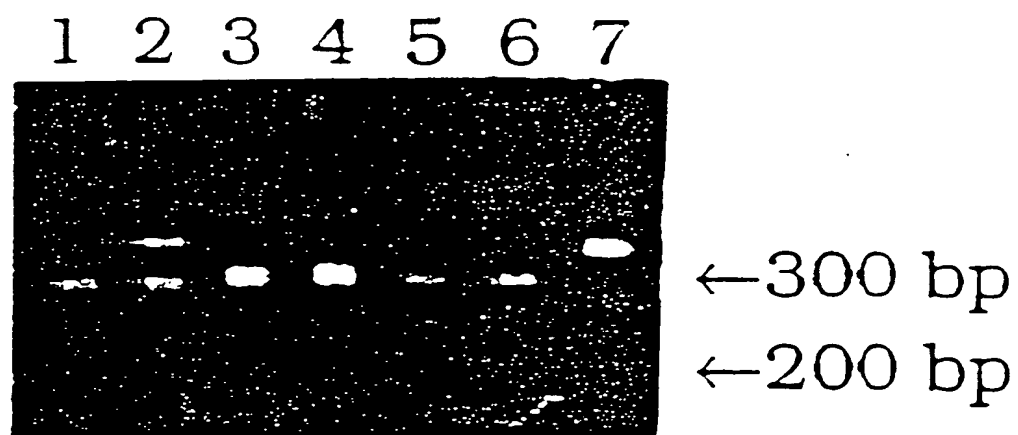
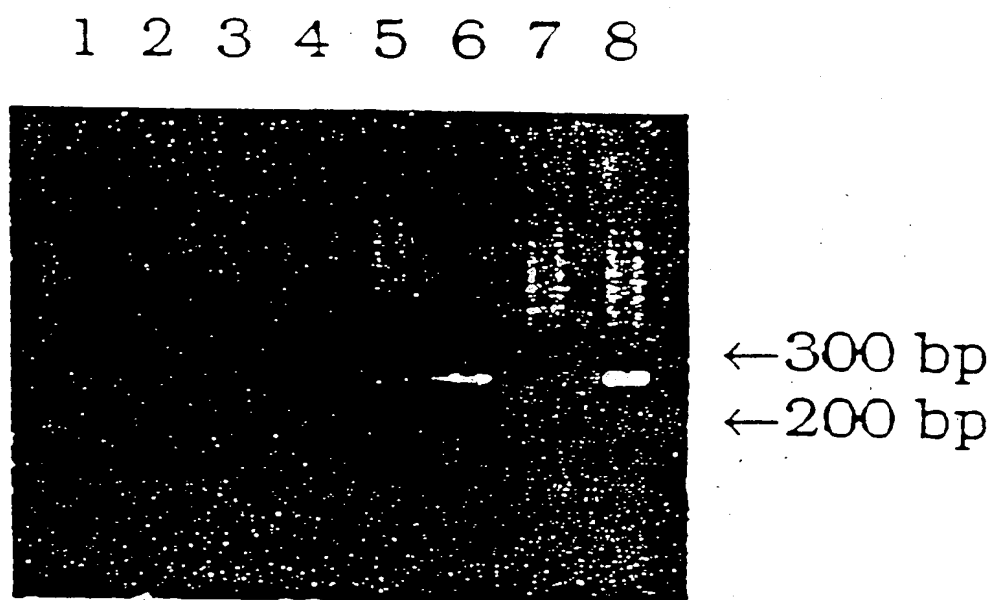


FIG. 3

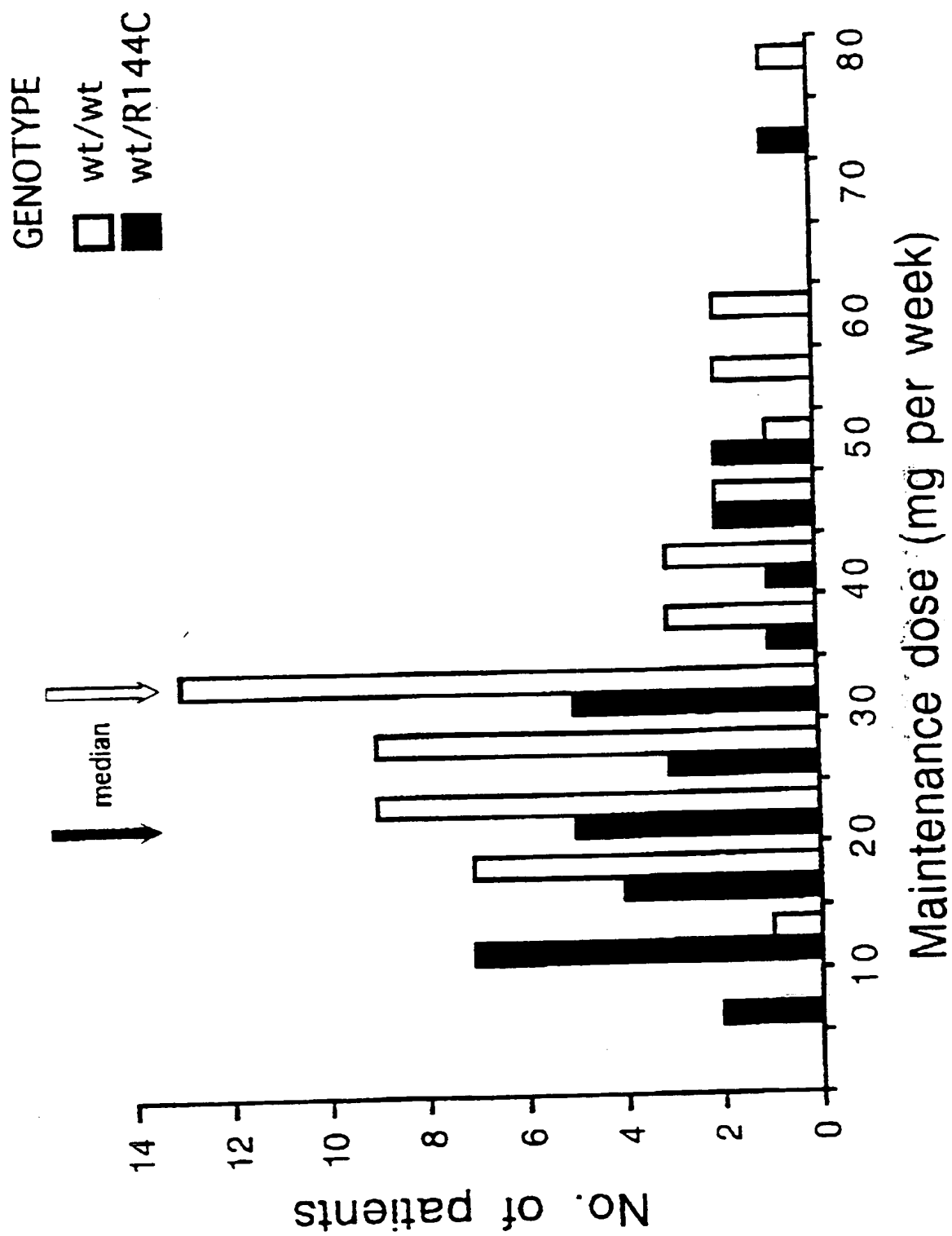
1 3 / 2 9



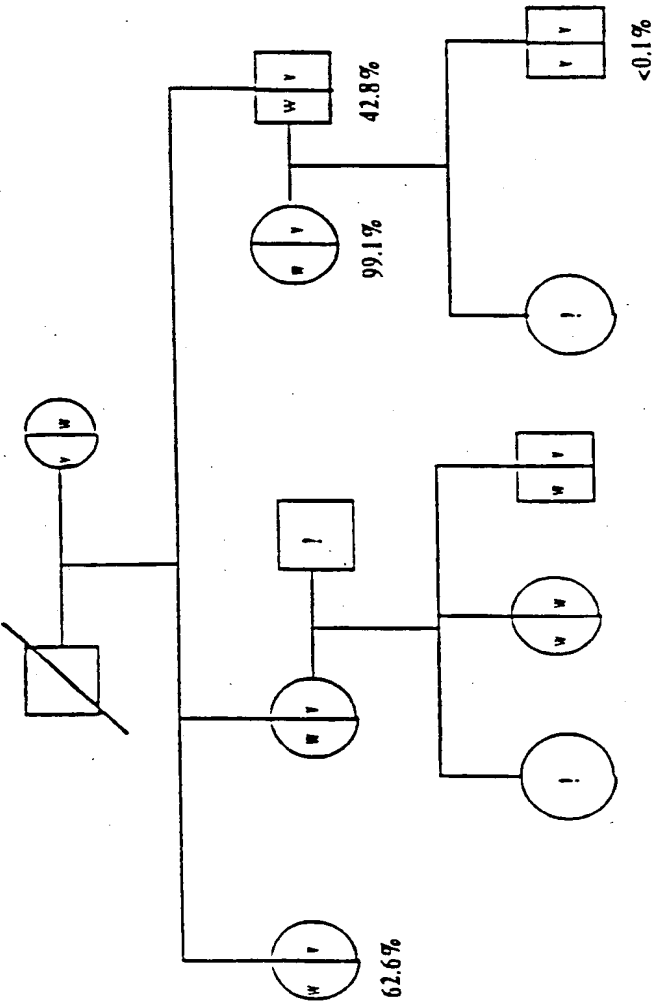
**FIG. 4**

1 4 / 2 9

FIG. 5



7-Hydroxylation of coumarin (%) in a family genotyped for the presence of *CYP2A6* and *CYP2A6v* alleles, showing subject homozygous for *CYP2A6v* who is deficient in coumarin 7-hydroxylation



w = *CYP2A6* wild-type  
v = *CYP2A6v* mutant allele  
? = not determined

**FIG. 6**

1 6 / 2 9

2A6 cDNA	GCGTGGTATTCAGCAACGGGGAGCGGCCAAGCAGCTCCGGCGCTTCTCCAT			
2A6 gene		T	TG	
2A6 cDNA	CGCCACCCTGCGGGGACTTCGGGGTGGGCAAGCGAGGCATCGAGGAGCGCATC			
2A6 gene		A	C	
2A6 cDNA	CAGGAGGAGCGGGGCTTCCTCATCGACGCCCCCTCCGGGGCACTGGC			
2A6 gene		T	G	A
			A	GCA

Comparison of CYP2A6 cDNA and genomic sequences for exon 3

**FIG. 7**

17 / 29

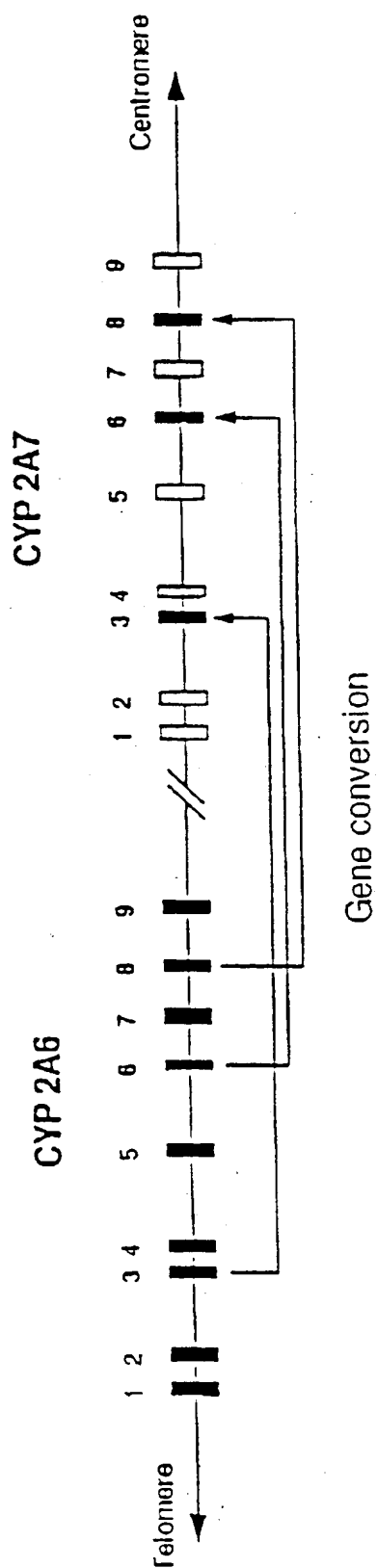


FIG. 8A

18 / 29

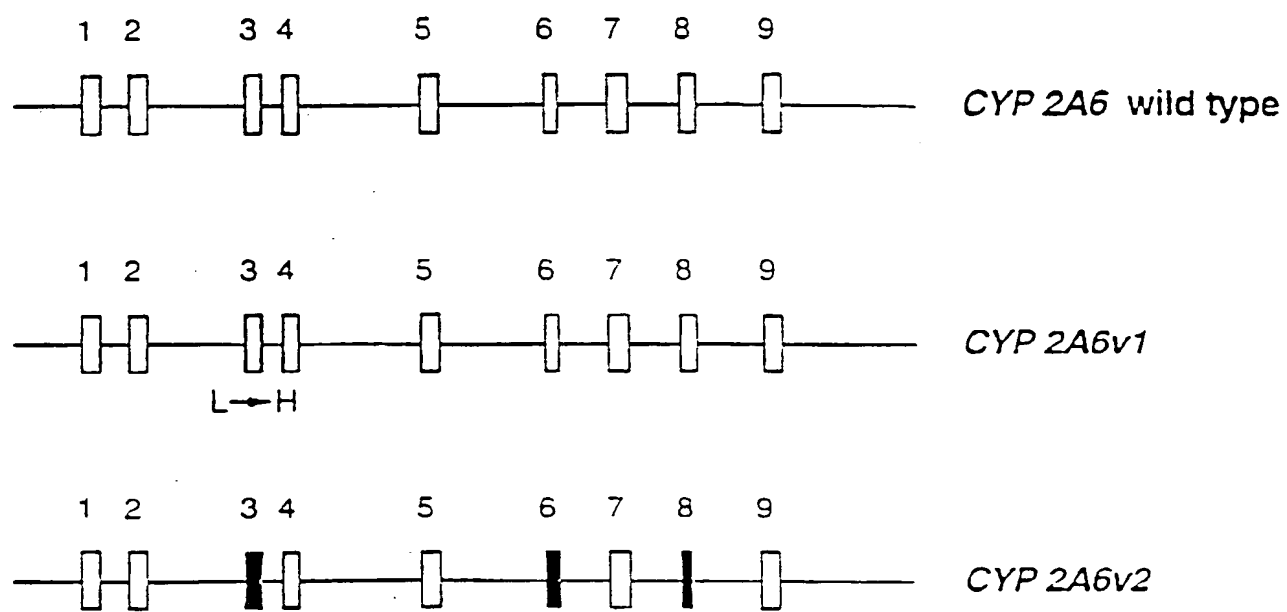


FIG. 8B

19/29

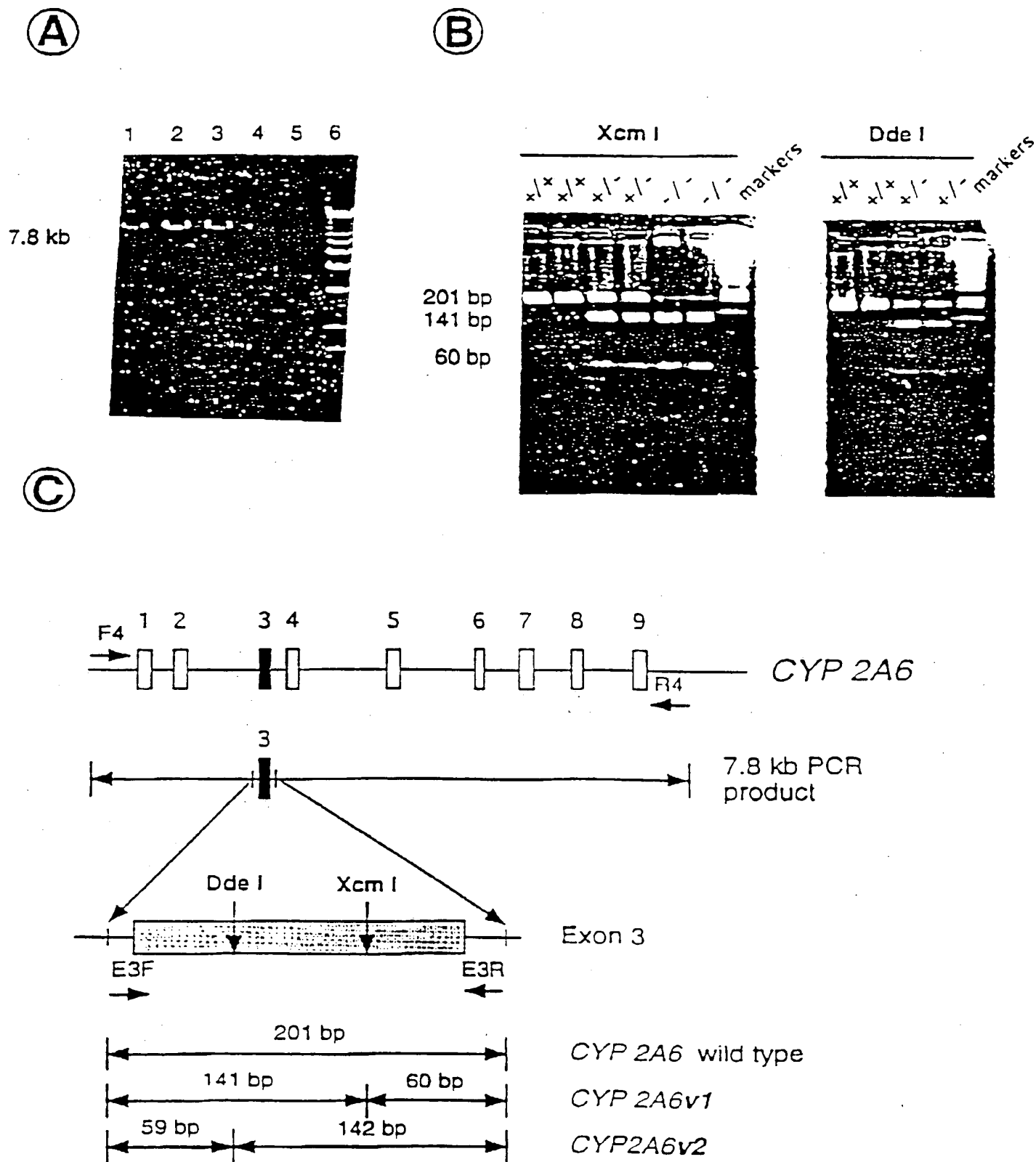


FIG. 9

20 / 29

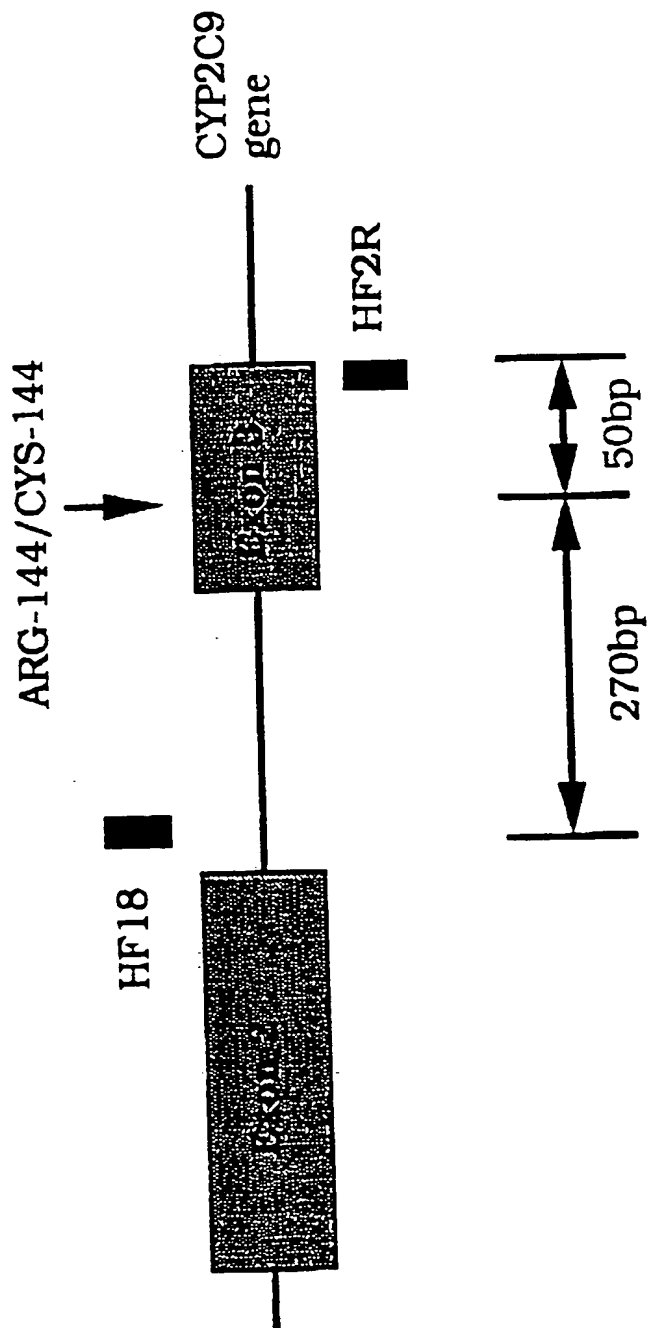


FIG. 10

21 / 29

## CYP2A6v2 cDNA.

5'TCTACCACCATGCTGGCCTCAGGGATGCTTCTGGTGGCCTTGCTGGCCTGCT  
GACTGTGATGGTCTTGATGTCTGTTTGGCAGCAGAGGAAGAGCAAGGGGAA  
GCTGCTCCGGGACCCACCCCATTTGCCCCCTCATTGGAAACTACCTGCAGCTGA  
ACACAGAGCAGATGTACAACCTCCCTCATGAAGATCAGTGAGCGCTATGGCC  
COGTGTTCAACATTCACTTGGGGGCCCCGGGGGTCGTGGTGTCTGTGTGGACATG  
ATGCOGTCAAGGGAGGCTCTGGTGGACCAGGCTGAGGAGTTCAGCGGGCGAGGC  
GAGCAAGCCACCTTCGACTGGGTCTTCAAAGGCTATGGCGTGGTATTCAGCA  
ACGGGGGAGCGGCGCCAAGCAGCTCCTGCGCTTTGOCATCGCCACCCCTGAGGGACT  
TOGGGGTGGGCAAGCGAGGCATCGAGGAGCGCATCCAGGAGGAGTCGGGCTTC  
CTCATCGAGGOCATCCGGAGCAOCCGCGGCCAATATCGATCCCACTTCTTC  
CTGAGCCGCACAGTCTCCAATGTATCAGCTCCATTGTCTTTGGGGACCGCTT  
TGACTATAAGGACAAAGAGTTCCTGTCACTGTTGCGCATGATGCTAGGAAT  
CTTCCAGTTCAAGTCAACCTCCACGGGGCAGCTCTATGAGATGTTCTCTTOGG  
TGATGAAACACCTGCCAGGACCACAGCAACAGGCCTTTCAGTTGCTGCAAGG  
GCTGGAGGACTTCATAGCCAAGAAGGTGGAGCACAACCAGCGCACGCTGGA  
TCCCAATTCCCCACGGGACTTCATTGACTCCTTTCTCATCCGCATGCAGGAGG  
AGGAGAAGAACCCCAACACGGAGTTCTACTTGAAGAACCTGATGATGAGC  
ACGTTGAACCTCTTCATTGCAGGCCACCGAGACGGTCAGCACCAACCTGCACTA  
TGGCTTCTTGCTGCTCATGAAGCAOCCAGAGGTGGAGGCCAAGGTCCATGAG  
GAGATTGACAGAGTGATCGGCAAGAACCGGCAGCCCAAGTTTGAGGACCGG  
GCCAAGATGCCCTACATGGAGGCAGTGATCCACGAGATCCAAAGATTGGA  
GACGTGATCCCATGAGTTTGGCCCCGAGAGTCAAAAAGGACACCAAGTTTC  
GGGATTTCTTCCTCCCTAAGGGCATAGAAGTGTTCCCTATGTTGGGCTCCGTG  
CTGAGAGACCTCAGGTTCTTCTCCAACCCCCGGGACTTCAATCCCCAGCACTTC  
CTGGGTGAGAAGGGGCAGTTTAAGAAGCGTGATGCTTTTGTGCCCTTCTCCA  
TCAGAAAGCGGAACCTGTTTCGGAGAAGGCCTGGCCAGAATGGAGCTCTTTCT  
CTTCTTCACCACCGTCATGCAGAACTTCGGCCTCAAGTCTCCAGTCACTA  
AGGACATTGACGTGTCCOCCAAACACGTGGGCTTTGCCACGATCCACGAAA  
CTACACCATGAGCTTCTGCCCCGCTGAGGAGGGCTGTGCOGGTGAAGGTCTG  
GTGGGCGGGGCCAGGGAAAGGGCAGGGCCAAGACCGGGCTTGGGAGAGGGGC  
GCAGCTAAGACTGGGGGCAGGATGGCGGAAAGGAAGGGGCGTGGTGGCTAG  
AGGGAAGAGAAGAAACAGAAAGCGGCTCAGTTCACCTTGATAAGGTGCTTCC  
GAGCTGGGATGAGAGGAAGGAAACCCCTTACATTATGCTATGAAGAGTAGT  
AATAATAGCAGCTCTTATTTCCTGA 3'

FIG. 11

22 / 29

1 AAGTTCCCTT GAAATATGGC TCTGGTCTTC CTCCCCTTGC CAATGAAGAA GATGGCAGTG  
61 GAGGTTCTAT GGCAGCCATC CTGGCCTCAC TCTGAGGTTT CAATGAGGAT TCTGGGCATC  
121 AAGAGACAGC TCTGGGCAAA GCTAAATCAA GTCAGCCCCC GGACCCAGTG CTGGGCTGCT  
181 GGGCTTTCTG GGAGAACGCC GCTGGGCTTG CTACACACTC CTCCTCCCAG AAACCTCCACA  
241 CCCACAGCCC TGGGTCTTCC TAGCCCCGAG ACTTTCAAGT CCATATGCCT GGAATCCCCC  
301 TTCCTGAGAC CCTTAACCCT GCATCCTCCA CAACAGAAGA CCCCTAAATG CACAGCCACA  
361 CTTTGTCTTA CCCTAATAAA ACCCAGACCT TTGGATTCCCT CTCCCCTGGA ACCCCAGAT  
421 CCGCACAAC TTTGGGTGCA TTCTCACTCT CAGACCCCAA ATCCAAAGCC CAAGTGTCTC  
481 CCTATGCAAA TATTCCAAAC TCCTCAGTTC TACAGCTTAT CTGTGCCCCC CTCCTAAATC  
541 CACAGCCCTG CGGCACCCCT CCTGAAGTAC CACAGATTTA GTCTGGAGGC CCCCTCTCTG  
601 TTCAGCTGCC CTGGGGTCCC CTTATCCTCC CTGTCTGGCT GTGTCCCAAG CTAGGCAGGA  
661 TTCATGGTGG GGCATGTAGT TGGGAGGTGA AATGAGGTAA TTATGTAATC AGCCAAAGTC  
721 CATCCCTCTT TTTCAAGCAG TATAAAGGCA AACCACCCCA GCCGTACCA TCTATCATCC  
781 CTCTACCACC ATGCTGGCCT CAGGGATGCT TCTGGTGGCC TTGCTGGCCT GCCTGACTGT  
841 GATGGTCTTC ATGTCTGTTT GGCAGCAGAG GAAGAGCAAG GGGAAAGCTGC CTCCGGGACC  
901 CACCCCATTG CCCTTCATTG GAAACTACCT GCAGCTGAAC ACAGAGCAGA TGTACAACCT  
961 CCTCATGAAG GTGTCCCAAG ACAGGGAGAT GGGTGTCTCG GGGTGGGGGC TGCCTAGTTG  
1021 GCTGGGGCTT TGTTGCCAGG GGTGACCAG TGTGGACCAG AGTCTTAGGA AATGGAGTTT  
1081 TGGAGTTTCA GCATCAGAAA GACAGGATCT TGGGATGTCC AGCTCCCTGA CTGTGAGAAC  
1141 CTGGGTGCGA AGCATCCCAG CACATGACAT CTCGGTGCTG GGCCCATTC AGAGTGGAGG  
1201 GTTCTCCCTC TAACCACTCC CACCCACCTC CATCAGATCA GTGAGCGCTA TGGCCCCGTG  
1261 TTCAACATTC ACTTGGGSCC CCGGCGGGTC GTGGTGCTGT GTGGACATGA TGCCGTCAGG  
1321 GAGGCTCTGG TGGACCAGGC TGAGGAGTTC AGCGGGCGAG GCGAGCAAGC CACCTTCGAC  
1381 TGGGTCTTCA AAGGCTATGG TGCCCAAGAG GGGGAAGGTG GCGAGGTGGA CACGAAGGTC  
1441 TCAGTGTTC CAGCCTTCTC CTTGACTCTC CTGACAACCT GAGGATAAGG GAGAGTCCCC  
1501 AGTCTGGTCT TCCCTCCCCA TCTCCTTACA TTGGGGCCTC TCCATGTGTA TCCCTACCT  
1561 GTCTCCAGCG GCCCTGTCTT GATTCCTCCC TGCCTCTCTC TGCCCACTC CTTATTCTC  
1621 TCTCACTGGA GTCTCCTCTT TCCCCTCTCT CTCCATCTCT AAGGACATCC TGGGTTTCTG  
1681 TTACCAGCC CTGGGTCTCT GTCTACATGA GTCTTTGAGG CCCTCTTAGC TTCTGGGCTT  
1741 CTCTGGGTTT CTCATCTCTC CGGATCCCTT TCTCAATTCT TCCTCTGTCT TAGGATGCCA  
1801 GGGTTATTCC TACTTCCACA TCTTCAGGCT CCATCTCTCT GTAACAGTCT CTCTTCTTC  
1861 CAGACCCCTCT CTGTTTCTAT CTCAATATTA AACTCTCTGC TCCAGCTCAG CTTAAGAATC  
1921 TCACACCAAG AGAGGATGTC CTCCACCCAG ATCTCCCCAT ATCTCACTAC CCCACCCCTC  
1981 ATCCTCTGCC TCCATCACTC TCTTTCTCTC CCCACTGCNC CTGCGGACGC GATCCAATGG  
2041 AGTGTGGAGC TAATGCCGTG AAGCTATGTG CATCTCTCTG TCTGGCCGTA CCTGGGTAAT  
2101 AACCTGATCG ACTAGGCGTG GTATTTCAGCA ACGGGGAGCG CGCCAAGCAG CTCTTGGCT  
2161 TTGCCATCGC CACCTTGAGG GACTTCGGGG TGGGCAAGCG AGGCATCGAG GAGCGCATCC  
2221 AGGAGGAGTC GGGCTTCTCT ATCGAGGCCA GGAACCAACC CAGGACGAGG AACC CGCGG  
2281 CGAGTGGGGG GGCAGGAGAA GGAAACCAACC CAGGACGAGG AACC CGCGG  
2341 GGGGATGGGG ACTAGGTGGG GAAAGGCGCC CGCACTTCCA GCCCTGGAGT CTGGCGCTGG  
2401 GAATTTGGCT CAACAAGGCC CTGCTCTCTG GAATTCTGAC TCTCTCAGA CCTCTGAGTT  
2461 GACTCTCTCC CCAACCCCTT TCTCCCGACA TACCCGGAGG CGCCAATATC GATCCCACCT  
2521 TCTTCTGAG CCGCACAGTC TCCAATGTCA TCAGCTCCAT TGTCTTTGGG GACCGCTTTG  
2581 ACTATAAGGA CAAAGAGTTC CTGTCACTGT TGCGCATGAT GCTAGGAATC TTCCAGTTCA  
2641 CGTCAACCTC CACGGGGCAG GTAATGGTTG CAGCCCGGCC CGTGAAGGCC CTTACCAAAA  
2701 CCGGCAAATT GTTCCCTTAC CCGGGGAAGG GGGCCCCAAA TTCCACCGC CCCCCGACA  
2761 GTGTCCCTC AAAATCAGTC CCCGATTGG GCAAATTGGC AGAGTGAAC CAGACCCGGG  
2821 TTGGTTGTCC AATCCCTTGC TCTCCAGGGA CACCGGGATA GCACAACAGA TGCTCCCCAA  
2881 AACAGAGCCT GCTGGCAGGA TGCATACCTT CAGCTCAGCT CTCTCACCTT GGGCACGTGT  
2941 TCCCATCCCC AACTTACCGG TAATTTCTAA CAGATGCTCC CTACCCAGGT CTTCTTGAAT  
3001 ATTTTAACAC CCGGAAACCC TGGGTACCTA ACCTTCCCTG TAAACTTTAG AGATTAAGTT  
3061 CTATCCGGCC CCTCTGAAAT ACCTAACCAC CGGAGACCAG ATGCCCTTAA CTCAGTTCTT  
3121 TCCTTGCTAT GAAACAAATC CCATTCCCAT CAGCTCCTGC CCCGTGACAG CTGTCTTCC  
3181 CTTCCCATCC TCTCTCTGCA ACCCCAGCTC TATGAGATGT TCTCTTCGGT GATGAAACAC  
3241 CTGCCAGGAC CGCAGCAACA GGCCTTTCAG TTGCTGCAAG GGCTGGAGGA CTTTCATAGCC  
3301 AAGAAGGTGG AGCACAACCA GCGCAGCTG GATCCCAATT CCCCACGGGA CTTTCATTGAC  
3361 TCCTTTCTCA TCCGATGCA GGAGGTACAC CCCAGCAGCC ACTCGGGGGA GATGCAAGC

FIG. 12 (Sheet 1)

23 / 29

3421 CAGGCAGAGG GAAATCAGTC TGGGAGTGGG GCAGGCAGAT GACACAGGCC CATTCAAATT  
3481 AACCTTCATC ATAATAATCC TCACAATTGG CTGGGTGCCG TGGCTAACAG CCTGTAATCC  
3541 CAGCACTTTG GGAGGCCGAG GCAGGTGGAT CACCTGAGGT CAGGAGTTGG AGACCAGCCT  
3601 GGCCAACATG GTCAAACCCC GTCTCTACTA AAAATCCAAA AATTAGTTGG GCATGGTGGC  
3661 GCGAAGGGGG GCAGAGGTTG CAATGAGCCA AGATCACGGC ATTGCACTCC AGTCTGGGTG  
3721 ACAGAATGAG GCCCTGTGTC AAAAAAATT AATCACTTGT TAAAAAGTA AGTGAGCCTG  
3781 CATGGTCATG CGCATGTGCA GCTCCAGCTA CTCAGGAGGC TGAGGCTGGA GGATTGCTTG  
3841 AGCTCAGGAG TTGGCGTCCG GCCTGTGCAA CTTAGCAAGA CCAAGTCAGT ATAAGAAAAA  
3901 AAAAAACAA AAAAAAGCT GACAGCTAAG TTGATAATTG ACGGACAGAT GGTGAGCAAG  
3961 GTAACGAAGG TGAGAAGGAA GAGCATTTGG GGCAACGCCA GGAGTCAGGG CAAGGGCTGG  
4021 TTCTTAGACC GAGTCTGGTA GGATCTAGGG CCCCTCTTCT CCACCCCTCG GTCTTGCCCC  
4081 AAAGAGAGGT CGAGGGTGCT GGGATTGCGC TAGACTCGAG TCTGTGTAGA TCTTGGGGTC  
4141 CCCTCTTGAC CCCCATTTGG CTGAACCTAA GAGTGGAAGA TCCATGGGGT GAACCCCTAG  
4201 ATGGTGCCCT GAGGTCAAGC AGGAGTGAGG TTGCTCTAAA GCCCCCTCTC CTTTCAGGAG  
4261 GAGAAGAACC CCAACACGGA GTTCTACTTG AAGAACCTGA TGATGAGCAC GTTGAACCTC  
4321 TTCATTGCGG GCACCGAGAC GGTCAGCACC ACCCTGCACT ATGGCTTCTT ACTGCTCATG  
4381 AAGCACCCAG AGGTGGAGGG TAAGGCTGGA GGGGGACGGA AGTGGAGGGC CCCAGACCCT  
4441 CAAAATTCCC CTTGCACTGG TGCAATGTCC CCACCTGTCC CAGATCCCGG GACCCTGAGA  
4501 CGTGACTTGC TGTCCAGAGA CAGGGCAACA TTCAGCTGGT AGGCATCAGC TGACTTCTG  
4561 TAGATATTAA AATATTGAAA ATGTCTGCAC TGATTGGTCA GTCACCTCTG TCCCAAGCCC  
4621 ACTGAGTGCC CACTGCCCCG TCCACCGGGT CATCCCCTAA GTTCTCTCCC GTGCTCTCCC  
4681 TGTGATTCTG GCACAACCTG GTTAACAGGA TCCTACTCCA ACAATGCGAA TGGGTGATGT  
4741 CTGTTCTGTT ATGAATGCTC TACTTCCGTC TCATAGGCGG AGGCATTTCA TCCACCCCAT  
4801 TTGTCCTATC CGGACTATCA TTTCTGCTC TGAGACCCCT AGATACCTAA ACACATTCCC  
4861 CCTCCTCCCC CAGCCAAGGT CCATGAGGAG ATTGACAGAG TGATCGGCAA GAACCGGCAG  
4921 CCCAAGTTTG AGGACCGGGC CAAGATGCCC TACATGGAGG CAGTGATCCA CGAGATCCAA  
4981 AGATTGAGG ACGTGATCCC CATGAGTTTG GCCCCGAGAG TCAAAAAGGA CACCAAGTTT  
5041 CGGGATTTCT TCCTCCCTAA GGTGCTATCC GCCCCACCC CCCAGACTAC GGGCACTCCA  
5101 GCCCCCTCTT GTGTCCCCAG CATCCCACCC ACATTAGAAG CCGTTCCACC TTTCCACTTA  
5161 TCCCTCAATC AGTCAAAAAA GACTTCCCCA ACCACCACAT CTTTCTAGAC CCGTCTCCAC  
5221 GACACTCCTG AGTCTTCGAT CTCTCCAGAC TCTTTGTGTC AGGAGAATCA AACACATGTT  
5281 CCCAAACTTC CTATCTTAAG AAACAGAAGC CCCCCTTCCA TTCGGCCTTT TGTCTAGGG  
5341 ACAGAAATCT CAGGTCCCCC AAACCTCTGC CTAGAAGGAC ATGGACCCCA TGTCTCCCAA  
5401 ACTTCTCTGT TCAGAGATGT GAACCTTCTA TCCCCAAGG TCCTCCCTCA GAGGTCCCCA  
5461 ATTCCCCTGC CTGCCACTTC CCTCACCCGG GGCACCCCTAG TTCCCCCTCC AGCCCCCTGT  
5521 TACTCTCAAC AATCCCCCAA CCCGCTCAT CACATACACC TTCTCTCTCC CTCGCCGGC  
5581 ATAGAAGTGT TCCCTATGTT GGGCTCCGTG CTGAGAGACC TCAGGTTCTT CTCCAACCCC  
5641 CGGGACTTCA ATCCCCAGCA CTTCCTGGGT GAGAAGGGGC AGTTTAAGAA GCCTGATGCT  
5701 TTTGTGCCCT TCTCCATCAG TAAGAGACCA CTGTTTGGTG CCAGGCTTAC TACTCACACC  
5761 AGCAGGGGCC TCCCTTACCC AGTCCCCCTC TCTGCCGTGT AGCCCTAGTAT TTCCCCAGCT  
5821 TGGCAAGTTC CTGTTAGCAA TCTACCGTCG AGCCACCAGG TGATACTCCC TTAACCTACCA  
5881 AGCACCCAGT ACCTGTGCCC AGGCAAAAGG AAAGGAAACA TCATACCCCT TTCAGAGGCG  
5941 GGGGAAAACC AAAGGCCAGA GAGAATCAGA GATTTATTTT CCTAGGGTCA CACAGGAGAT  
6001 TCTTCAGCAT CCCTAAAAAG GAGATGACGG CACAGCAGGT CATATTTGGG AGTCTTATC  
6061 TGGGGGAAGG GGGATCTTAA ACCTCCCAT GTGGACACCT GGCATCGATC AACCCTATCT  
6121 TTTGGTCTAT TTTTGGGTCA CTCAAGGAAA CTGAGGTCAA GGAGGGTCAA GAGGCTCCCT  
6181 CTTAAAGTCT CTCAGGGCCA TATATTCCAC CCTTCCCTCC TGGGAGAGCC GCAGCTGGAG  
6241 GTCGGTACTG GGGCGAGGCT GCACTGAGAG TGGGCTTCAC CTCCACCCCT CCCGCTCTC  
6301 CTCCTCAGGA AAGCGGAAGT GTTTCGGAGA AGGCCTGGCC AGAATGGAGC TCTTTCTCTT  
6361 CTTACACCAC GTCATGCAGA ACTTCCGCTT CAAGTCCCTC CAGTCACCTA AGGACATTGA  
6421 CGTGTCCCCC AAACACGTGG GCTTGTCCAC GATCCCACGA AACTACACCA TGAGCTTCTT  
6481 GCCCCGCTGA GCGAGGGCTG TCCCGGTGAA GGTCTGGTGG GCGGGGCCA GGAAGGGCA  
6541 GGGCCAAGAC CGGGCTTGGG AGAGGGGGCC AGCTAAGACT GGGGGCAGGA TGGCGGAAAG  
6601 GAAGGGGGCT GGTGGCTAGA GGAAGAGAA GAAACAGAAG CAGGCTCAGT CACCTTGATA  
6661 AGGTGCTTCC GAGCTGGGAT GAGAGGAAG AAACCCCTAC ATTATGCTAT GAAGAGTAGT  
6721 AATAATAGCA GCTCTTATTT CCTGAGCAGG TACCCCGTG TCACCTTTGT TCAAAAACCA  
6781 TTGCACGCTC ACCTAATTTG CCACAAAACC CCCTTCGAAG GGGGCTTCAT GCCCATTTTA

FIG. 12 (Sheet 2)

24 / 29

6841 CACGTGACAA AACTGAGGCT TAGAAAGTTG TCTCTGATGT CTCACAAAAC ATAAGTGCCC  
6901 AGAAAAATCTG CGAACACAGA TCTGTGCCCCA TAGCCTTCTA GACAGATTCT TAAAAAGCAC  
6961 CTATTCTCTCA CGCAAAACAG TTTAGTATAG AATCACATGG CCTGAACATC CCTGTCCGGG  
7021 GGAGTTCCCC AGAGACCTGG GGGGTGGTTG CCCTGCCCTC ACTGCACACA TGCCCACT  
7081 CTCACCTACT CAACATGCTG TGAATACCCG GGTGTAATCT GTGCTTGCTA CCAGATAAGG  
7141 CCACTGTAGC CCATTCAGAG TCAGCCCAGG GACACAACGA GACATCACTG GACATACAGG  
7201 GTCAGTCCAT TAACAA

25 / 29

## CYP2A13 cDNA

5' ATGGCCACCATGCTGGCCTCAGGGCTGCTTCTGGTGACCTTGCTGGCCTGCTT  
GACTGTGATGGTCTTGATGTTCAGTCTGGCGGCAGAGGAAGAGCAGGGGGGAA  
GCTGCTCCGGGAOCCACCCCATTTGCCCTTCATTGGAAACTACCTOCAGCTGAA  
CACAGAGCAGATGTACAACCTCCCTCATGAAGATCAGTGAGCGCTATGGCCT  
GTGTTCAOCCATTCAGTTGGGGCCCCGGGGGTGGTGGTGGTGGTGGGACATGAT  
GCGTCAAGGAGGCTCTGGTGGACAGGCTGAGGAGTTCAGCGGGGAGGCGA  
GCAGGCCAOCCTTCGACTGGCTCTTCAAAGGCTATGGCGTGGCGTTCAGCAACG  
GGGAGCGCGCCAAGCAGCTCCGGCGCTTCTCCATCGCCACCCCTAAGGGGTTTTG  
GCGTGGGCAAGCGGGCATCGAGGAACGCATCCAGGAGGAGGCGGGCTTCTC  
ATCGACGCTCCGGGGCAGCAGCGCGCCAATATCGATCCCAOCTTCTTCTG  
AGCCGCACAGTCTCCAATGTTCATCAGCTCCATTGTCTTTGGGGACCGCTTTGA  
CTATGAGGACAAAGAGTTCCCTGTCACTGTTGCGCATGATGCTGGGAAGGTTC  
CAGTTCACGGGAACCTCCACGGGGCAGCTCTATGAGATGTTCTCTTCGGTGAT  
GAAACACCTGCCAGGACCACAGCAACAGGCCTTTAAGGAGCTGCAAGGGCT  
GGAGGACTTCATCGCCAAGAAGGTGGAGCACAACCAGCGCAOCTGGATCCC  
AATTCOCCACGGGACTTCATCGACTCCTTTCTCATCCGCATGCAGGAGGAGGA  
GAAGAACCCCAACACAGAGTTCTACTTGAAGAACCTGGTGATGAOCCACCT  
GAOCTCTTCTTTGOGGGCACTGAGACCGTGAGCAOCCACCTGCGCTACGGTTT  
CCTGCTGCTCATGAAGCAOCCAGAGGTGGAGGCCAAGGTCCATGAGGAGATT  
GACAGAGTGATCGGCAAGAACCGGCAGOCCAAGTTTGAGGAOCCGGGCOCAAG  
ATGCCCTACACAGAGGCAGTGATCCACGAGATCCAAAGATTTGGAGACATG  
CTOCCCATGGGTTTGGCCACAGGGTCAACAAGGACACCAAGTTTCCGGGATT  
TCTTCCCTCCCTAAGGGCACTGAAGTGTTCCCTATGCTGGGCTCCGAGCTGAGA  
GACCCAGGTTCTTCTCCAACCCOCCAGGACTGCAGTCCOCCAGCACTTCTGGAT  
GAGAAGGGGGCAGTTTAAGAAGAGTGATGCTTTTGTGCCCTTTTCCATCGGA  
AAGCGGTACTGTTTTGGAGAAGGCCTGGCCAGAATGGAGCTCTTTCTCTTCT  
TCAOCCACCATCATGCAGAACTTTGGCTTCAAGTCCCCTCAGTCGCCTAAGGAT  
ATCGACGTGTCCOCCAAACACGTGGGCTTTGCCACGATCCCAOCAAACCTACAC  
CATGAGCTTCTGCCCCGCTGAGCGAGGGCTGTGCTGGTGCAGGGCTGGTGGGC  
GGGGCCAGGGAAAAGGGGGGGGAGGGGGGGGCTTTGGGGAGGGGGGGGGCT  
AAGAATGGGGGCGAGTGGGGGAAGGAAGGGGAGAGGTGGTTAGAGGGGAACA  
GAAGAAACAGAAGGGGCTCAGTTCACCTTGATGATGTCCTTCAGAGCTGTG  
ATGAGAGGAAGGGAAACCTTACAGTATGCTACAAAGAGTAGTAATAATA  
GCAGCTCTTATCTCCTGA 3'

FIG. 13

26 / 29

3421 CCTCCACTTC AGCATCTTCA CCAGCCCCAC TTTATACCTG AGCACCTGAA CAAAAGCCCC  
3481 CAATCCAGAC CCAGTAAGTA TCTGGACAGC TGTCTCCAAC CAAGTCCACT TGAATGCCTA  
3541 AATACCTAGA CAGGTGCCAC TCACCTCATA CCAGCCCCAC CTGAAGAGCT AAACACCTGG  
3601 ACAGCTGTCT TCCAACCTCA CTTCACTTGA ATATCTGAAC ACCTAGATGT GTGCTCCAAT  
3661 CCAGCCTCAT TTGCATACCT GAAACCTGGA TATATGCCTC AGTTCTTCTC ACCTAAATTA  
3721 CTAGACCGTG CCCCTGGCAC CTAATCCACG TGAAAACCTA GATATAAGTT TCCATCCAAC  
3781 CCCACTGAAA TACCTAAACA CCTGGACAGA TGCCTTTAAC TCCGTTCCCTT CCTTGCTATG  
3841 AAACAAATCC CCATTCCCCT CAGCTCCTGC CCCGTGACAG CTGTCTTCC CTTCCTCATC  
3901 TCTCTCTGCA ACCCCAGCTC TATGAGATGT TCTCTTCGGT GATGAAACAC CTGCCAGGAC  
3961 CACAGCAACA GGCCTTTAAG GAGCTGCAAG GGCTGGAGGA CTTCATCGCC AAGAAGGTGG  
4021 AGCACAACCA GCGCACGCTG GATCCCAATT CCCACGGA CTTCATCGAC TCCTTTCTCA  
4081 TCCGCATGCA GGAGGTACAT CCCAGCAGCC AGTGCAGGCA GGTGCAAAGC CAGGGAGAGG  
4141 GAAATCAGGA TGGGAGTGGG GTGGGCAGAC GACACAGGCC CATTCAAATT AGCCCTCGTC  
4201 ATAATAATCC TTACAATTGG CCAGGCGCGG TGGCTCATGA CCTGTAATCC CAGCACTTTG  
4261 GGAGGCCGAG GCAGGTGGAT CACCTGAGGT CAGGAGTTCC AGACCAGCCT GGCCAACATG  
4321 GTGAAACCCC GTCTCTACTA AAAATACAAA AATGAGCTAG GTATGGTGGC ATGCGCCTGT  
4381 AATCCCAGCT ACTCAGGAGG CTGAGACAGA AGAATTTGTT TGAATCCGGG AGGCAGAGGT  
4441 TGCAGTGAGC CGGGATCATG CCACTGCACCT CCGGCCTGAG TGACAGAGCA AGACCTGTGA  
4501 AAAAAAAAAA AAAAAAAAAA AAAAAATTCC GGAAAACCCC AATTACATCA CCCACTGTG  
4561 TCCCATCTAC TGAGCCCTCA CCCACAAGGA CGGCTTATGG AGGTGGATTGA GATTGGAAG  
4621 AACTTCTCAA GAACCTACCG GTGCCAGGAA CTGGGTAAAG TGTTTTATGA TAGTCCGCCA  
4681 TGAACACTT TTAACAGTTC TTGAGGGAGG TTCCTCATG GCCCCAGTTG TACAAATGAG  
4741 GAAACTGAGG CCCAGAGAGT TTAAGTGTCT TAACTGAGGT CACAACAGTG AGGAAGACCA  
4801 TGGTCCCCCT AGCTCAAACC CTGGTCTCTC TGAGCCTATA CCTGGTGTCT TTAGCCACCA  
4861 TGCTCTCTAA CCGTTCATGT CCTGGTTAGC AGACACACCT CTGTGGACAG GTGACCTGGC  
4921 TTTACATTCG AGGGTCCCCG CCTACCTCTG GATGTCAGCC TCCCATGTGG GAAGGCTTTA  
4981 GGAAGCCAAA GCTCAGGGAG AAAGGATCAA GGGAGGGATT CCTCCACAGT AAGTTTCAAG  
5041 ATTTTTAGGG AAGAAATAGG ATGCTGTGTC TTAATAATTCT GTGCTTGTAT CTCAGAAAAA  
5101 CTCTTTTTTT CTGACTCTTC ATCTTGCCAT CTCTGTACTA CTTCTCTTTC GTCTCCCCCTC  
5161 ATCCTTCTCT TTCCAAATAT TCCTATCATT AAAAAAGTAA CAGACTGGGA AACATGGCAA  
5221 AACCCCGTCT GTACAAAAAA ATGGCTAGGC ATGGTGGTGC ATGCCTGCGG TCCAGCTAC  
5281 TAAGGAGGTT GAGGTGGGAG GATATCTTGA GCCCAGGGTG GGCAGAGCTT TCAATGAGCC  
5341 GATATCACAG CCCTGCCCTC CAGCCTGGGT GACAGAATAA GACCGTGTCT CCAAAAAA  
5401 AAAAGAATTA ATTTTTTAAC AGTTAACAAG TGAGCCTGCA TAGTCATGTG CATGTGCCAGT  
5461 TCCAGCTACT CTGGAGGCTG AGACCGGAGG ATTCCTTGAA CCCAGGAGTT GGAGTCCAGC  
5521 CTGTGCAACT TAGCAAGACC AAGTCTGCAT AAAAAAATAA AAAACCAACT GACAGCTAAG  
5581 TTGACAATTA AAGGATAGAT GATCAGTGAG GTAAAGAAGG TGAGAAGGAA GAGCATTTTG  
5641 GGCAAAGCCA GCAGCCAGGG CAAGGGCTGG AACCTGGAGC GAGTTTGGCA AATCTAGGGT  
5701 CCCTCTTTCC ACCTTTGGTC TGGACCAAAG AGAGGTAGCT CCAAAGGAAA AGCCCTAGAA  
5761 GGGCCCCAAG AGCATGGAGA GTGAGCTTGG TCTAAACCGC CCTCTCCCTG CAGGAGGAGA  
5821 AGAACCCCAA CACAGAGTTC TACTTGAAGA ACCTGGTGAT GACCACCTG AACCTCTTCT  
5881 TTGCCCGCAC TGAGACCGTG AGCACCACCC TGCGCTACGG TTTCTGTCTG CTCATGAAGC  
5941 ACCCAGAGGT GGAGGGTAAG ACTGGAAAGG GAGGAAAGTG AAGGGCCCCA GACCCTCAAA  
6001 ACTCCCCCTGA GCCTGGTGCA GTGTACCCAC CTATCCCAGA TCCCAGGACC CTGAGACGTG  
6061 CCTTGCTGTG CAGAGACAGG ACAATATTCA GCTGATAGGC ATCAGCTGAG TCTCATTAGC  
6121 TATTAAAAATA TTGAAAATGT CTGCACATGAT TGGTCAGTCA CTCCTGTCCC AAGCCCACTG  
6181 AGTGTCCGCT GCCTGCTCCT CTGGATCATC CCCTAAGTTC CTCCTTTGTC CTACCCTGTG  
6241 ATTCTGACAC AACCTGGTTT AACAGGGATC CTGCTGCAAA CAATGCGAAT GGGTGATGTC  
6301 TTGTTCTTGT TTATGAATGG GCTTACCCTT CGTGTGAGAG GTGGAAGCTA TGTCAACCGC  
6361 CGTGTTTTAG CTAGGGGGGG CGATACATGC CCTGCTCTAA GACCCCTAGA GAGGGTAAAG  
6421 ATATTCCCCCT CCTCCGCCAG CCAAGGTCCA TGAGGAGATT GACAGAGTGA TCGGCAAGAA  
6481 CCGGCAGCCC AAGTTTGAGG ACCGGGCCAA GATGCCCTAC ACAGAGGCAG TGATCCACGA  
6541 GATCCAAAGA TTTGGAGACA TGCTCCCCAT GGGTTTGGCC CACAGGGTCA ACAAGGACAC  
6601 CAAGTTTCGG GATTTCTTCC TCCCTAAGGT GCTGTCTCCC CTCCACCACC ACCACTCAGA  
6661 CTACGGGGAC TTCCAGCCTC TCTCTGTGTC CCCAGAATCC TGCCCCCATT AGTGTCTTAG  
6721 ACTCTGTCCC ACTCCCTCAA TCAGTCAAAA AAGACTTCCC CAACCACCAC ATCTGTTCCTA  
6781 CCTTTCCACT TAGACAGTCC TGAGTCTGTC ATCTCGCCAG ACTCTTTGTG TCAGGAGAAT

FIG. 14 (Sheet 1)

27/29

```

3421 CCTCCACTTC AGCATCTTCA CCAGCCCCAC TTTATACCTG AGCACCTGAA CAAAAGCCCC
3481 CAATCCAGAC CCAGTAAGTA TCTGGACAGC TGTCTCCAAC CAAGTCCACT TGAATGCCTA
3541 AATACCTAGA CAGGTGCCAC TCACCTCATA CCAGCCCCAC CTGAAGAGCT AAACACCTGG
3601 ACAGGTGTCT TCCAAC TCACTTGA ATATCTGAAC ACCTAGATGT GTGCTCCAAT
3661 CCAGCCTCAT TTGCATACCT GAAACCTGGA TATATGCCTC AGTTCTTCTC ACCTAAATTA
3721 CTAGACCGTG CCCCTGGCAC CTAATCCACG TGAAAAC TTA GATATAAGTT TCCATCCAAC
3781 CCCACTGAAA TACCTAAACA CCTGGACAGA TGCCTTTAAC TCCGTTCCTT CTTGCTATG
3841 AAACAAATCC CCATTCCCAT CAGCTCCTGC CCCGTGACAG CTGTCTTCC CTTCCCATCC
3901 TCTCTCTGCA ACCCCAGCTC TATGAGATGT TCTCTTCGGT GATGAAACAC CTGCCAGGAC
3961 CACAGCAACA GGCCTTTAAG GAGCTGCAAG GCCTGGAGGA CTTTCATCGCC AAGAAGGTGG
4021 AGCACAACCA GCGCACGCTG GATCCCAATT CCCACGGGA CTTTCATCGAC TCCTTTCTCA
4081 TCCGCATGCA GGAGGTACAT CCCAGCAGCC AGTGCAGGCA GGTGCAAAGC CAGGGAGAGG
4141 GAAATCAGGA TGGGAGTGGG GTGGGCAGAC GACACAGGCC CATTCAAATT AGCCCTCGTC
4201 ATAATAATCC TTACAATTGG CCAGGCGCGG TGGCTCATGA CCTGTAATCC CAGCACTTTG
4261 GGAGGCCGAG GCAGGTGGAT CACCTGAGGT CAGGAGTTCC AGACCAGCCT GGCCAAACATG
4321 GTGAAACCCC GTCTCTACTA AAAATACAAA AATGAGCTAG GTATGGTGGC ATGCGCCTGT
4381 AATCCCAGCT ACTCAGGAGG CTGAGACAGA AGAATTTGTT TGAATCCGGG AGGCAGAGGT
4441 TGCAGTGAGC CGGGATCATG CCACTGCACT CCGGCCTGAG TGACAGAGCA AGACCCTGTA
4501 AAAAAAAAAA AAAAAAAAAA AAAAAATTC GGAAAACCCC AATTACATCA CCCACTGCTG
4561 TCCCATCTAC TGAGCCCTCA CCCACAAGGA CGGGTTATGG AGGTGGATTA GATTGGAAG
4621 AACTTCTCAA GAACTACCGG GTGCCAGGAA CTGGGTTAAG TGTTTTATGA TAGTCCGCCA
4681 TGAACACTT TTAACAGTTC TTGAGGGAGG TTAAGTTGCT TAAGTGGGT CACAACAGTG AGGAAGACCA
4741 GAAACTGAGG CCCAGAGAGT TTAAGTTGCT TAAGTGGGT CACAACAGTG AGGAAGACCA
4801 TGGTCCCCCT AGCTCAAACC CTGGTCTCTC TGAGCCTATA GCTGGTGCTT TTAGCCACCA
4861 TGCTCTCTAA CCGTTCATGT CCTGGTTAGC AGACACACCT CTGTGGACAG GTGACCTGGC
4921 TTTACATTGC AGGGTCCCCG CCTACCTCTG GATGTCAGCC TCCCATGTGG GAAGGCTTTA
4981 GGAAGCCAAA GCTCAGGGAG AAAGGATCAA GGGAGGGATT CCTCCACAGT AAGTTTCAAG
5041 ATTTTTAGGG AAGAAATAGG ATGCTGTTGC TTAAATTTCT GTGCTTGAT CTCAGAAAAA
5101 CTCTTTTTTT CTGACTCTTC ATCTTGCCAT CTCTGTAATA CTTTCTCTTC GTCTCCCCTC
5161 ATCTTCTCT TCCAAATAT TCCTATCATT AAAAAAGTAA CAGACTGGGA AACATGGCAA
5221 AACCCCGTCT GTACAAAAAA ATGGCTAGGC ATGGTGGTGC ATGCCTGCGG TCCAGCTAC
5281 TAAGGAGGTT GAGGTGGGAG GATATCTTGA GCCCAGGGTG GGCAGAGGTT TCAATGAGCC
5341 GATATCACAG CCCTGCCCTC CAGCCTGGGT GACAGAATAA GACCGTGTCT CCAAAAAAA
5401 AAAAGAATTA ATTTTFTAAC AGTTAACAAG TGAGCCTGCA TAGTCATGTG CATGTGCAGT
5461 TCCAGCTACT CTGGACGCTG AGACCGGAGG ATTCTTGAA CCCAGGAGTT GGAGTCCAGC
5521 CTGTGCAACT TAGCAAGACC AAGTCTGCAT AAAAAAAAAA AAAACCAACT GACAGCTAAG
5581 TTGACAATTA AAGGATAGAT GATCAGTGAG GTAAAGAAGG TGAGAAGGAA GAGCATTTTG
5641 GGCAAAGCCA GCAGCCAGGC CAAGGCTGG AACCTGGAGC GAGTTTGGCA AATCTAGGGT
5701 CCGCTTTTCC ACCTTTGGTC TGGACCAAAG AGAGGTAGCT CCAAAGGAAA AGCCCTAGAA
5761 GGGCCCCAAG AGCATGGAGA GTGAGCTTGG TCTAAACCGC CCTCTCCCTG CAGGAGGAGA
5821 AGAACCCCAA CACAGAGTTC TACTTGAAGA ACCTGGTGAT GACCACCCTG AACCTCTTCT
5881 TTGCGGGCAC TGAGACCGTG AGCACCACCC TGCCCTACGG TTTCTGTCTG CTCATGAAGC
5941 ACCCAGAGGT GGAGGGTAAG ACTGGAAAGG GAGGAAAGTG AAGGGCCCCA GACCCTCAAA
6001 ACTCCCTGTA GCCTGGTGCA GTGTACCCAC CTATCCCAAG TCCAGGACC CTGAGACGTG
6061 CCTTGCTGTC CAGAGACAGG ACAATATTCA GCTGATAGGC ATCAGCTGAG TCTCATTAGC
6121 TATTAAAATA TTGAAAATGT CTGCACTGAT TGGTCAGTCA CTCCTGTCCC AAGCCCCATG
6181 AGTGTCCGCT GCCTGCTCCT CTGGATCATC CCCTAAGTTC CTCCCTTGTC CTACCCTGTG
6241 ATTCTGACAC AACCTGGTTT AACAGGGATC CTGCTGCAAA CAATGCGAAT GGGTGATGTC
6301 TTGTTCTTGT TTATGAATGG GCTTACCCTT CGTGTGAGAG GTGGAAGCTA TCTCAACCGC
6361 CGTGTTTTAG CTAGGGGGGG CGATACATGC CCTGCTCTAA GACCCCTAGA GAGGGTAAAG
6421 ATATTCCCTT CCTCCGCCAG CCAAGGTCCA TGAGGAGATT GACAGAGTGA TCGGCAAGAA
6481 CCGGCAGCCC AAGTTTGAGG ACCGGGCCAA GATGCCCTAC ACAGAGGCAG TGATCCACGA
6541 GATCCAAAGA TTTGGAGACA TGCTCCCCAT GGGTTTGGCC CACAGGGTCA ACAAGGACAC
6601 CAAGTTTCGG GATTTCTTCC TCCCTAAGST GCTGTCTCCC CTCCACCACC ACCACTCAGA
6661 CTACGGGGAC TTCCAGCTC TCTCTGTGTC CCCAGAATCC TGCCCCCATT AGTGTCTTAG
6721 ACTCTGTCCC ACTCCCTCAA TCAGTCAAAA AAGACTTCCC CAACCACCAC ATCTGTTCCT
6781 CCTTTCCTCT TAGACAGTCC TGAGTCTGTC ATCTCGCCAG ACTCTTTGTG TCAGGAGAAT

```

FIG. 14 (Sheet 2)

28 / 29

6841 ACACCCCATG TTCCCAATCT TCCTGTCTTA AGAAACAGAA GCCCCCTTC CATTAGGCCT  
6901 TGTGGCTTAG GGACACAAAT CTCAGGTCCC TCAAACACCC TGGCTAGTGG AACATGGACC  
6961 CCATGTCTCC CAAACTTCCT GTCTCAGAGA CATGAACTT CTATCCCCCA AAGCTCCTCC  
7021 CTCAGAGGTC CCCAACTCCT CCATGTCTGT CCACTCCCCG CACCTGGGGG ACCCTAGAGC  
7081 CCCCTGGAGC CCCTGTGTAC TTTCACCAAT CCCCCCAACC TGGCTCATAA CACACACCTT  
7141 CCTCCTCCCT CCCAGGGCAC TGAAGTGTTC CCTATGCTGG GCTCCGAGCT GAGAGACCCC  
7201 AGGTTCTTCT CCAACCCCCA GGAAGTGCAGT CCCCAGCACT TCCTGGATGA GAAGGGGCAG  
7261 TTAAAGAAGA GTGATGCTTT TGTGCCCTTT TCCATCGGTA AGAGACACTG TTTGCTGCCA  
7321 GGCCACGGCT CACACCAGCA GGGGCCCTCT CACCCACCT CCCCCTCTCT CGGTGTAGCC  
7381 TGGTATTTCT CCAGCTTGA AGTTCTCTGT AGAATCTACC ATTGAGCCGC CACCAGCTGA  
7441 TACTCCCTTA ACTGCCAAGC ACCCAATACC TGCGCCCGAG TAAAGGGGAA GGAAACATCT  
7501 TCCCCCATAG ATTTATTTGT CTAGGGTCAC ACAGCAGATT CTTGAGCTCC CTGAAAAGGA  
7561 GATAATGGTA CAGCACAGCA GTCATATTTG CAAGTGTATC TGGGGGGTAG GGGCATCTAA  
7621 ACCTCCCAT TGTACACCTG GCATGGATCA CCCCATCTAT GATGGAGGCA TGACATTATG  
7681 CCTTTTTCGA AACCCTATGA ACTGTATAAC ACAGAGTAAA CCCTAATGTA AACTATGGAC  
7741 TTTGGTTAGT AATAATATAT CAATATTGGT TCACCATTGT TATATCTCTT ATAGAAGGAA  
7801 ACTGAAGCTC AGGGAGGATC GGAGTCTCCT CTGAAAGTCT CTCAGGCCAT AATATTCCCA  
7861 CCCCTCCTCC CTAGAGAGTG CAGCCGGGGG TCAGTAGGGG TTGAGGCTGC ACTGAGAGTG  
7921 GGCTTCACCT TCACCCCTCC TGCTCTCTCT CCTCAGGAAA GCGGTACTGT TTTGGAGAAG  
7981 GCCTGGCCAG AATGGAGCTC TTTCTCTTCT TCACCACCAT CATGCAGAAC TTTGCTTCA  
8041 AGTCCCTTCA GTCCCTAAG GATATCGACG TGTCCCCCAA ACACGTGGGC TTTGCCACGA  
8101 TCCCACGAAA CTACACCATG AGCTTCCTGC CCCGCTGAGC GAGGGCTGTG CTGGTGCAGG  
8161 GCTGGTGGC GGGGCCAGGG AAACGGCCGG GGCAGGGGGG GGGCTTGTGG GAGGGGCCGG  
8221 GCTAAGAATG GGGGCAGTGG GGGGAAGGAAG GGGAGAGGTG GTTAGAGGGA ACAGAAGAAA  
8281 CAGAAGGGGC TCAGTTCACC TTGATGATGT CCTTCAGAGC TGTGATGAGA GGAAGGGAAA  
8341 CCTTACAGTA TGCTACAAAG AGTAGTAATA ATAGCAGCTC TTATCTCTG AACAAAGTCC  
8401 TCCCTGTCAG CTTTGTTCAA AAAGCGTTGC ACGCTCACCT CACTTATTTG CCACACACCT  
8461 CTACCAATGG GGGAAAAGTC TTCATTCCCC TTTTACACG TGAGAAAGGT GCGGCTCAGA  
8521 AAGTTGTCTC TATCTGAAAA CTCACAAAC GCAAGTGTCC AGAGGATCTT GGAACACAGA  
8581 TCTGGGCCCC TAGCCCTCTA GATCGATCCT CACCATAGCA CCCCCTCTTC ACGTAAAATA  
8641 GCTTAGTATA GCATCACATG GCCTGAACAC CCCTGGGCGG GGGGGTTCCC CAGAGACCTG  
8701 GCGGGCGGCT GCCCTGCCTA CTCTGTACAC TCGCCTACTC GGGACGATCC GGGCACCAGG  
8761 GTGTCACCTG AGCTCGCTA

FIG. 14 (Sheet 3)

2 9 / 2 9

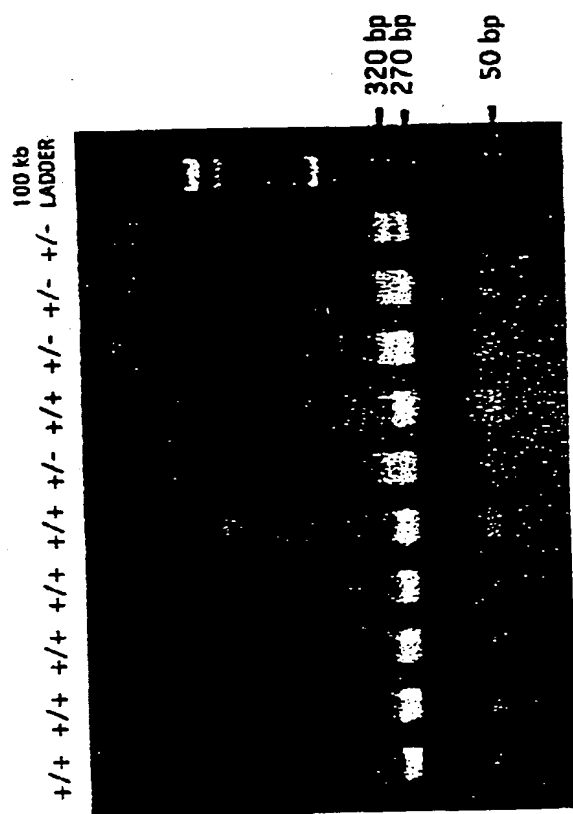


FIG. 15

**PCT**WORLD INTELLECTUAL PROPERTY  
ORGANIZATION  
International

INTERNATIONAL APPLICATION PUBLISHED UNDER

WO 9534679A3

(51) International Patent Classification<sup>6</sup> :**C12N 1/68, 9/02, C12P 19/34, A61K  
31/70, C07H 21/04, C12N 15/55****A3**

(11) International Publication Number:

**WO 95/34679**

(43) International Publication Date: 21 December 1995 (21.12.95)

(21) International Application Number: PCT/US95/07605

(22) International Filing Date: 16 June 1995 (16.06.95)

(30) Priority Data:

9412054.0	16 June 1994 (16.06.94)	GB
9502728.0	13 February 1995 (13.02.95)	GB
9507640.2	12 April 1995 (12.04.95)	GB

(71) Applicant (for all designated States except US): THE GOVERNMENT OF THE UNITED STATES OF AMERICA, represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Office of Technology Transfer, National Institutes of Health, Suite 325, 6011 Executive Boulevard, Rockville, MD 20852 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): GONZALEZ, Frank, J. [US/US]; Unit 101, 5000 Battery Lane, Bethesda, MD 20814 (US). IDLE, Jeffrey, R. [GB/GB]; 19 Leazes Crescent, Newcastle, Upon Tyne NE1 4LW (GB).

(74) Agents: FEILER, William, S. et al.; Morgan & Finnegan, L.L.P., 345 Park Avenue, New York, NY 10154 (US).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

**Published***With international search report.**Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(88) Date of publication of the international search report:

21 March 1996 (21.03.96)

(54) Title: DEFECTS IN DRUG METABOLISM

**(57) Abstract**

The invention relates to genetic material, and specifically portions of DNA, for identifying the presence or absence of a mutation in the drug metabolism gene CYP2C9 and CYP2A6. Further, the invention comprises a method for determining such mutations and a kit incorporating the genetic material of the invention for performing the said methods so as to determine the presence or absence of mutations in the drug metabolizing gene CYP2C9 and CYP2A6.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

# INTERNATIONAL SEARCH REPORT

In: Application No  
PCT/US 95/07605

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 C12Q1/68 C12N9/02 C12P19/34 A61K31/70 C07H21/04  
C12N15/55

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6 C12Q C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	AMERICAN JOURNAL OF HUMAN GENETICS, vol. 57, September 1995 pages 651-660, FERNANDEZ-SALUERO, P. ET AL 'A genetic polymorphism in coumarin 7-hydroxylation' see the whole document ---	1-6, 8-13, 17, 20
X	WO,A,91 01990 (CITY OF HOPE) 21 February 1991 see page 5 ---	5,6
X	WO,A,94 01548 (MEDICAL RESEARCH COUNCIL) 20 January 1994 see page 364; claim 1 ---	5,6
-/--		

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

\*&\* document member of the same patent family

Date of the actual completion of the international search

30 January 1996

Date of mailing of the international search report

2. 02 96

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+ 31-70) 340-3016

Authorized officer

Osborne, H

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CARCINOGENESIS, vol. 14 , no. 7, July 1993 pages 1421-1427, TIANO, H. ET AL 'retroviral mediated expression of human cytochrome P450 2A6 in CH3/10T1/2 cells confers transformability by NNK.' see the whole document ----	1-4
A	BIOCHEMISTRY, vol. 29, 1990 pages 1322-29, YAMANO S ET AL 'The CYP3A3 gene product catalyses coumarin hydroxylation in human liver microsomes' see the whole document see page 1324, line 1; figure 1 -----	1

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 95/07605

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 26-28  
because they relate to subject matter not required to be searched by this Authority, namely:  
Remark: Although claims 26-28 as far as they concern an "in vivo" method, are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 95/07605

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9101990	21-02-91	US-A- 5081235	14-01-92
		AU-B- 629401	01-10-92
		AU-B- 6344790	11-03-91
		CA-A- 2035899	27-01-91
		DE-D- 69018801	24-05-95
		DE-T- 69018801	19-10-95
		EP-A- 0436016	10-07-91
		JP-T- 4502258	23-04-92
		US-A- 5075431	24-12-91
WO-A-9401548	20-01-94	AU-B- 4512193	31-01-94
		EP-A- 0587279	16-03-94